4 - DESIGUALDAD DE CHEBYSHEV- LEY DE LOS GRANDES NUMEROS

La desigualdad de Chebyshev es una importante herramienta teórica. Entre otras aplicaciones constituirá un medio para comprender cómo la varianza mide la variabilidad de una dada variable aleatoria, con respecto a su esperanza matemática. También nos permitirá establecer con más precisión el hecho, reiteradamente señalando, de que la frecuencia relativa f_A de un suceso A asociado a un experimento aleatorio ε tiende, cuando el número de repeticiones de ε se hace infinitamente grande, a la probabilidad $P(A)$ (resultado conocido como la Ley de los grandes números). Pero además es de utilidad práctica pues, al constituir una cota de ciertas probabilidades, nos podrá servir como una estimación de esas mismas probabilidades.

4.1-Desigualdad de Chebyshev

Sea X una variable aleatoria cuya esperanza es $E(X)$, sea c un número real cualquiera y supongamos que $E\left[(X-c)^2\right]$ existe y es finito. Entonces

$$\boxed{\forall \varepsilon > 0 \qquad P(|X-c| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} E\left[(X-c)^2\right]}$$

Dem.) Consideraremos el caso en que X es una v.a. continua. El caso de una v.a. discreta se demuestra en forma similar cambiando integrales por sumas.

Sea, entonces, $f(x)$ la *fdp* de X. Tenemos:

$$P(|X-c| \geq \varepsilon) = \int_{x:|x-c|\geq\varepsilon} f(x)dx$$

Ahora bien, los valores de x que verifican $|x-c| \geq \varepsilon$ son los mismos que verifican $\dfrac{(x-c)^2}{\varepsilon^2} \geq 1$. Entonces, puesto que $f(x) \geq 0$ y que ambos miembros en la desigualdad anterior son también no negativos es:

$$P(|X-c| \geq \varepsilon) = \int_{x:|x-c|\geq\varepsilon} 1.f(x)dx \leq \int_{x:|x-c|\geq\varepsilon} \frac{(x-c)^2}{\varepsilon^2}.f(x)dx \leq \int_{-\infty}^{\infty} \frac{(x-c)^2}{\varepsilon^2}.f(x)dx$$

donde la última desigualdad proviene del hecho de que simplemente extiendo los límites de integración a todos los reales pero siendo siempre el integrando positivo. De manera que estoy agregando una contribución positiva sobre el valor de la integral $\displaystyle\int_{x:|x-c|\geq\varepsilon} \frac{(x-c)^2}{\varepsilon^2}.f(x)dx$.

Si tenemos presente la expresión para la esperanza de una función $H(X)$ de una variable aleatoria X:

$\int_{-\infty}^{\infty} H(x)f(x)dx = E[H(X)]$ y lo aplicamos a $H(x)=\dfrac{(x-c)^2}{\varepsilon^2}$, tenemos finalmente:

$$P(|X-c| \geq \varepsilon) \leq E\left[\frac{(X-c)^2}{\varepsilon^2}\right] = \frac{1}{\varepsilon^2}E\left[(X-c)^2\right]$$

<u>Observación:</u> La desigualdad lleva el nombre del matemático ruso que la descubrió. Su nombre aparece en una variedad de formas en la literatura: Chebyshev, Chebychev, Tchebyshev, etc.

<u>Formas alternativas de la desigualdad de Chebyshev.</u>

Podemos escribir la desigualdad de Chebyshev en una serie de formas alternativas:

a_1) Tenemos en primer lugar la forma que acabamos de demostrar:

$$P(|X-c| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}E\left[(X-c)^2\right]$$

a_2) Si consideramos el suceso complementario a $|X-c| \geq \varepsilon$, es decir $|X-c| < \varepsilon$, podemos escribir, recordando que $P(\overline{A}) = 1 - P(A)$:
$P(|X-c| < \varepsilon) = 1 - P(|X-c| \geq \varepsilon) \geq 1 - \dfrac{1}{\varepsilon^2}E\left[(X-c)^2\right]$, esto es:

$$P(|X-c| < \varepsilon) \geq 1 - \frac{1}{\varepsilon^2}E\left[(X-c)^2\right]$$

b_1) Si en a_1) elegimos $c = E(X)$ tenemos $P(|X-E(X)| \geq \varepsilon) \leq \dfrac{1}{\varepsilon^2}E\left([X-E(X)]^2\right)$ y recordando la definición de la varianza: $V(X) = E\left([X-E(X)]^2\right)$ tenemos:

$$P(|X-E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

b_2) La correspondiente expresión para el suceso complementario es:

$$P(|X-E(X)| < \varepsilon) \geq 1 - \frac{V(X)}{\varepsilon^2}$$

c_1) Si en a_1) elegimos $c = E(X)$ y además elegimos $\varepsilon = k\sigma_X = k\sqrt{V(X)}$, tenemos $P(|X-E(X)| \geq k\sigma_X) \leq \dfrac{V(X)}{(k\sigma_X)^2} = \dfrac{\sigma_X^2}{(k\sigma_X)^2}$, es decir:

$$P(|X - E(X)| \geq k\sigma_X) \leq \frac{1}{k^2}$$

c_2) Finalmente a correspondiente expresión para el suceso complementario en c_1) es:

$$P(|X - E(X)| < k\sigma_X) \geq 1 - \frac{1}{k^2}$$

A continuación daremos un ejemplo en el que podremos apreciar cómo la desigualdad de Chebyshev nos permite tener estimaciones de ciertas probabilidades que en algunos casos mejoran las "estimaciones" triviales dadas por el axioma i) de las probabilidades, esto es, que $0 \leq P(A) \leq 1$.

<u>Ejemplo</u> Deseamos estimar la probabilidad $P\left(|X - E(X)| \geq \frac{3}{2}\sigma_X\right)$

a) Sin conocer la distribución
b) Sabiendo que $X \sim U\left[1 - \frac{1}{\sqrt{3}},\ 1 + \frac{1}{\sqrt{3}}\right]$.

a) Una estimación de la probabilidad en consideración cuando no se conoce la distribución, es decir que vale cualquiera sea la distribución puede tenerse usando la desigualdad de Chebyshev en la forma c_1) : $P(|X - E(X)| \geq k\sigma_X) \leq \frac{1}{k^2}$ con $k = \frac{3}{2}$.
Tenemos en consecuencia:

$P\left(|X - E(X)| \geq \frac{3}{2}\sigma_X\right) \leq \frac{1}{(3/2)^2} = \frac{4}{9} = 0.44$. Vemos que, en este caso, la estimación mejora sustancialmente la cota superior trivial 1:

$$P\left(|X - E(X)| \geq \frac{3}{2}\sigma_X\right) \leq 0.44 < 1$$

Notemos que esta estimación es aplicable cualquiera se la distribución de X y vale sea ésta discreta o continua.

b)Si sabemos que $X \sim U\left[1 - \frac{1}{\sqrt{3}},\ 1 + \frac{1}{\sqrt{3}}\right]$ tenemos toda la información y podemos calcular la probabilidad exactamente.

Tenemos:

$$E(X) = \frac{\left(1 + \frac{1}{\sqrt{3}}\right) + \left(1` - \frac{1}{\sqrt{3}}\right)}{2} = 1$$

$$\sigma_X = \sqrt{V(X)} = \sqrt{\frac{\left[\left(1+\frac{1}{\sqrt{3}}\right)-\left(1`-\frac{1}{\sqrt{3}}\right)\right]^2}{12}} = \sqrt{\frac{\left(2/\sqrt{3}\right)^2}{12}} = \sqrt{\frac{4}{36}} = \sqrt{\frac{1}{9}} = \frac{1}{3}$$. Entonces:

$$P\left(|X-E(X)| \geq \frac{3}{2}\sigma_X\right) = P\left(|X-1| \geq \frac{1}{2}\right) = 1-P\left(|X-1| < \frac{1}{2}\right) = 1-P\left(\frac{1}{2} \leq X \leq \frac{3}{2}\right)$$

$$P\left(|X-E(X)| \geq \frac{3}{2}\sigma_X\right) = 1-\frac{\frac{3}{2}-\frac{1}{2}}{\left(1+\frac{1}{\sqrt{3}}\right)-\left(1-\frac{1}{\sqrt{3}}\right)} = 1-\frac{1}{\frac{2}{\sqrt{3}}} = 1-\frac{\sqrt{3}}{2} = 0.134$$. Este valor exacto es menor que la cota superior dada por la desigualdad de Chebyshev:

$$P\left(|X-E(X)| \geq \frac{3}{2}\sigma_X\right) = 0.134 \leq 0.44 < 1.$$

<u>La varianza como una medida de la concentración de la *fdp* de una v.a. alrededor de la esperanza.</u>

Podemos usar las formas *b*) de la desigualdad de Chebyshev para interpretar a la varianza $V(X)$ como una medida de la variabilidad de la variable aleatoria *X* con respecto a su esperanza o en otras palabras de cómo la distribución de la v.a. *X* se concentra o dispersa con respecto a la esperanza $E(X)$.

De la expresión $P(|X-E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$ vemos que, para un ε dado, si $V(X)$ es muy pequeño entonces la probabilidad de que *X* tome valores lejos de $E(X)$ es muy chica, es decir hay una gran probabilidad de que *X* tome valores próximos a $E(X)$. Inversamente si $V(X)$ es grande, la probabilidad de que *X* tome valores alejados de $E(X)$ puede ser también grande. Podemos precisar un poco más esto considerando el siguiente

Teorema. Si $V(X) = 0$ entonces $P[X = E(X)] = 1$. Decimos que $X = E(X)$ con probabilidad 1 (*X* es igual a su esperanza con probabilidad 1).

Dem.) Para cualquier $\varepsilon > 0$, si $V(X) = 0$ tenemos de b_2) :

$P(|X-E(X)| < \varepsilon) \geq 1-\frac{V(X)}{\varepsilon^2} \rightarrow P(|X-E(X)| < \varepsilon) = 1$, donde me quedo sólo con la igualdad porque la probabilidad no pude superar a 1.

Puesto que ε puede hacerse arbitrariamente pequeño, el teorema queda demostrado.

4.2 - La ley de los grandes números.

La ley de los grandes números establece en forma precisa el hecho que cuando el número de repeticiones de un experimento se hace muy grande, la frecuencia relativa f_A de un suceso *A* relacionado con el experimento converge en sentido probabilístico a la probabilidad $P(A)$. Daremos una versión de la Ley de los grandes números conocida como la ***forma de Bernoulli***.

Teorema (Forma de Bernoulli de la ley de los grandes números)
Sea un experimento probabilístico y sea *A* un suceso asociado con él. Consideremos *n* repeticiones independientes del experimento. Sea n_A el número de veces que ocurre *A* en las *n* repeticiones de forma tal que $f_A = \frac{n_A}{n}$ es la frecuencia relativa. Sea $P(A) = p$ (que se supone igual para todas las repeticiones). Entonces, para cualquier número $\varepsilon > 0$ se cumple

$$P(|f_A - p| < \varepsilon) \geq 1 - \frac{p(1-p)}{n\varepsilon^2}.$$

Dem.)
De acuerdo con su significado n_A, el número de veces que ocurre el suceso *A* en las *n* repeticiones, es una variable aleatoria distribuida binomialmente con parámetros *n* y *p*: $n_A \sim B(n, p)$. Luego:

$$\begin{cases} E(n_A) = np \\ V(n_A) = np(1-p) \end{cases}$$

Por lo tanto

$$\begin{cases} E(f_A) = E\left(\frac{n_A}{n}\right) = \frac{1}{n}E(n_A) = p \\ V(f_A) = V\left(\frac{n_A}{n}\right) = \frac{1}{n^2}V(n_A) = \frac{p(1-p)}{n} \end{cases}$$

En consecuencia, aplicando a la v.a. f_A la desigualdad de Chebyshev en la forma b_2)

$P(|X - E(X)| < \varepsilon) \geq 1 - \frac{V(X)}{\varepsilon^2}$, es decir,

$P(|f_A - E(f_A)| < \varepsilon) \geq 1 - \frac{V(f_A)}{\varepsilon^2}$ llegamos a lo propuesto por el teorema:

$$P(|f_A - p| < \varepsilon) \geq 1 - \frac{p(1-p)}{n\varepsilon^2}.$$

Es evidente que el *teorema* anterior implica que
$\lim_{n\to\infty} P(|f_A - p| < \varepsilon) = 1$ para todo $\varepsilon > 0$.

Entonces decimos que, en este sentido, la frecuencia relativa f_A converge a la probabilidad $P(A)$.

Observación: Esta convergencia, llamada ***convergencia en probabilidad*** difiere de la convergencia normalmente usada en Cálculo (límite aritmético).
Recordemos que una sucesión $\alpha_1, \alpha_2, ..., \alpha_n, ...$ tiene límite α o también que $\lim_{n \to \infty} \alpha_n = \alpha$ si $\forall \varepsilon > 0 \quad \exists \nu$ tal que $|\alpha_n - \alpha| < \varepsilon$ para todo $n > \nu$. Esto significa que, desde un n en adelante, α_n se aproxima permanentemente al valor límite α. En cambio cuando decimos que $f_A = n_A / n$ converge a $P(A)$, estamos significando que la probabilidad del suceso $\left\{ \, |f_A - P(A)| < \varepsilon \right\}$ puede hacerse arbitrariamente próximo a uno tomando un n suficientemente grande. Pero estamos hablando de probabilidad y no de certeza como en el caso del límite aritmético. Es decir, no significa que al tomar un *n* grande ocurra ciertamente que nos aproximemos más al valor de la probabilidad sino que existe una gran probabilidad de que eso ocurra.

5- VARIABLES ALEATORIAS BIDIMENSIONALES

5.1 – Generalidades

Hasta ahora hemos considerado el caso de variables aleatorias unidimensionales. Esto es, el resultado del experimento de interés se registra como un único número real.
En muchos casos, sin embargo, nos puede interesar asociar a cada resultado de un experimento aleatorio, dos o más características numéricas. Por ejemplo, de los remaches que salen de una línea de producción nos puede interesar el diámetro X y la longitud Y. Teniendo en cuenta la inevitable variabilidad en las dimensiones de los remaches debido a las numerosas causas presentes en el proceso de fabricación, los podemos representar asociándoles dos variables aleatorias X e Y que pueden pensarse como una ***variable aleatoria bidimensional***: (X,Y).

Sea ε un experimento aleatorio y S un espacio muestral asociado a él. Sean $X : S \to R$, $Y : S \to R$, que a cada resultado $s \in S$ le asignan el par de números reales (x, y)
Llamaremos a (X,Y) ***variable aleatoria bidimensional***.
Si en lugar de dos variables aleatorias, tenemos n variables aleatorias $X_1, X_2,..., X_n$, llamaremos a $(X_1, X_2,..., X_n)$ ***variable aleatoria n-dimensional***

En lo que sigue nos referiremos en particular a variables aleatorias n-dimensionales con n=2, es decir nos concentraremos en ***variables aleatorias bidimensionales*** por cuanto son las más simples de describir, fundamentalmente en relación a la notación. Pero debemos tener presente que las propiedades que estudiemos para ellas se pueden extender sin demasiada dificultad al caso general.

Al conjunto de valores que toma la variable aleatoria bidimensional (X,Y) lo llamaremos ***recorrido*** de la v.a. (X,Y) y lo indicaremos R_{XY}. En otras palabras $R_{XY} = \left\{ (x, y) : x = X(s) \ \ e \ \ y = Y(s) \ \ con \ \ s \in S \right\}$, es decir, es la imagen por (X,Y) del espacio muestral S.
Notar que el recorrido de (X,Y) es un subconjunto del espacio Euclidiano: $R_{XY} \subseteq R^2$. Como antes, puede considerarse al recorrido R_{XY} como un espacio muestral cuyos elementos son ahora pares de números reales.

Como con cualquier espacio muestral, según el número de elementos que lo constituyen, podemos clasificar a los recorridos R_{XY} en numerables (finitos o infinitos) y no-numerables.
Los recorridos numerables son, en general, de la forma

$$R_{XY} = \left\{ (x_i, y_j) \ \ con \ \ i = 1,2,...,n \ \ y \ \ j = 1,2,..m \right\} = \{(x_1, y_1),(x_1, y_2),...,(x_n, y_m)\} \qquad \text{(finito)}$$

$$R_{XY} = \left\{ (x_i, y_j) \ \ con \ \ i = 1,2,... \ \ y \ \ j = 1,2,.. \right\} = \{(x_1, y_1),(x_1, y_2),...\} \qquad \text{(infinito numerable)}$$

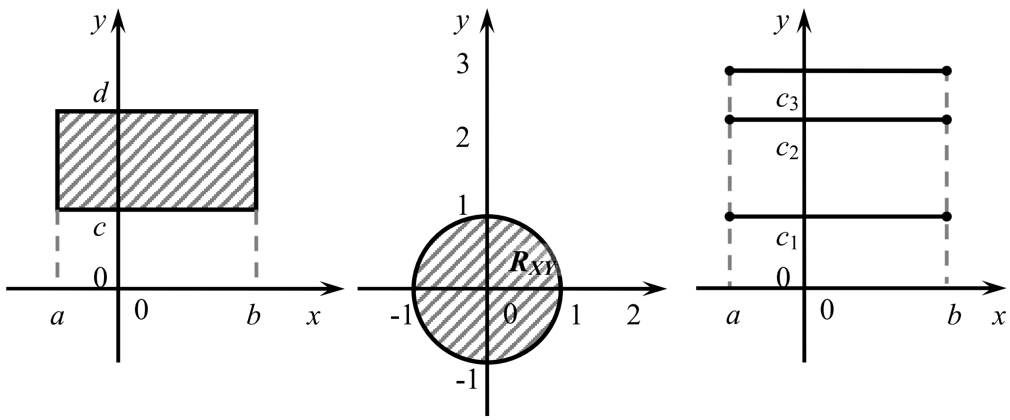
Los recorridos no numerables son regiones o subconjuntos no numerables del plano Euclidiano. Por ejemplo:

$$R_{XY} = \left\{ (x, y) : a \le x \le b; \ \ c \le y \le d \right\} \qquad \text{(no numerable)}$$

$R_{XY} = \{(x, y) : x^2 + y^2 \le 1\}$ (no numerable)

$R_{XY} = \left\{(x, y_j) : a \le x \le b,\ \ y_j = c_1, c_2, c_3\right\}$ (no numerable "mixto")

cuyas gráficas se pueden apreciar en la figura siguiente. Notar en el último recorrido, X es v.a. continua e Y discreta.



Clasificaremos a las variables aleatorias bidimensionales de la siguiente manera:
(X,Y) es v.a. ***bidimensional discreta*** si X e Y son discretas
(X,Y) es v.a. ***bidimensional continua*** si X e Y son continuas
El caso X continua, Y discreta (o viceversa) no lo consideramos.

Sea (X,Y) una variable aleatoria bidimensional discreta y sea R_{XY} su recorrido (numerable). Sea $p : R_{XY} \to R$ una función que a cada elemento (x_i, y_j) le asigna un número real $p(x_i, y_j)$ tal que $P\left(X = x_i,\ Y = y_j\right) = p(x_i, y_j)\ \ \forall (x_i, y_j) \in R_{XY}$ y que verifica.

a) $p(x_i, y_j) \ge 0 \quad \forall (x_i, y_j) \in R_{XY}$

b) $\sum_{(x_i, y_j) \in R_{XY}} p(x_i, y_j) = \sum_i \sum_j p(x_i, y_j) = 1$

A esta función la llamaremos ***función de probabilidad puntual conjunta*** de la variable aleatoria bidimensional (X,Y). En forma abreviada la designaremos *fdp conjunta*.

Ejemplos:

1-Dos líneas de producción, señaladas I y II, manufacturan cierto tipo de artículo a pequeña escala. Supóngase que la capacidad máxima de producción de la línea I es cinco artículos por día, mientras que para la línea II es 3 artículos/día. Debido a los innumerables factores presentes en todo proceso de producción, el número de artículos realmente producido por cada línea puede pensarse como una variable aleatoria. En conjunto podemos pensar en una variable aleatoria bidimensional (X,Y) discreta, donde la primera componente X corresponde a la producción de la línea I y la segunda componente Y a los artículos que salen de la línea II. La *fdp* conjunta correspondiente a variables aleatorias bidimensionales suele presentarse, por comodidad, como una tabla. Supongamos que la para la v.a. (X,Y) que nos interesa aquí la tabla correspondiente a $p(x_i, y_j)$ es

Y/X	0	1	2	3	4	5
0	0	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

¿Cuál es la probabilidad de qué salgan más artículos de la línea I que de la línea II?

Antes de calcular la probabilidad que nos pide el problema, hagamos algunas consideraciones sobre la tabla que representa a $p(x_i, y_j)$.
Se trata de una tabla a doble entrada donde en la primera fila se indican los valores que puede tomar la v.a. X (en este caso X=0,1,2,3,4,5) y la primera columna indica los valores que puede tomar la variable Y (0,1,2,3). Para determinar el valor de la $p(x_i, y_j)$ cuando la v.a. (X,Y) toma el valor (x_i, y_j) consideramos el número que se encuentra en la columna correspondiente a $X = x_i$ y la fila correspondiente a $Y = y_j$. Por ejemplo: $p(4,2) = P(X = 4, Y = 2) = 0.05$.
Podemos verificar fácilmente que la *fdp* conjunta definida por esta bien definida. En efecto verifica las condiciones a) $p(x_i, y_j) \geq 0 \quad \forall (x_i, y_j) \in R_{XY}$ y b) $\sum_{(x_i, y_j) \in R_{XY}} p(x_i, y_j) = 1$.

Para contestar la pregunta del enunciado, consideremos el suceso $B \subset R_{XY}$ definido

B: "es el suceso que ocurre cuando la línea I produce más artículos que la línea II" o, $B = \{X > Y\}$. Luego:

$$P(B) = P(X > Y) = \sum_{y_j=0}^{3} \sum_{x_i > y_j} p(x_i, y_j) = 0.01+0.03+0.05+0.07+0.09+0.04+0.05+0.06+0.08+$$
$$+0.05+0.05+0.06+0.06+0.05=0.75.$$

2- Hay tres cajas registradoras a la salida de un supermercado. Dos clientes llegan a las cajas en diferentes momentos cuando no hay otros clientes ante aquellas. Cada cliente escoge una caja al azar e independientemente del otro.
Sean las variables aleatorias X: " nº de clientes que escogen la caja 1" e Y: "nº de clientes que escogen la caja 2". Hallar la *fdp* conjunta de *(X,Y)*

Podemos suponer que el espacio muestral original S es el conjunto de pares ordenados $S = \{(1,1);(1,2);(1,3);(2,1);(2,2);(2,3);(3,1);(3,2);(3,3)\}$ donde la primera componente del par indica la caja elegida por el cliente 1 y la segunda componente del par indica la caja elegida por el cliente 2.
Además notar que X como Y pueden tomar los valores 0, 1, 2
El punto muestral (3,3) es el único punto muestral que corresponde al evento $\{X = 0, Y = 0\}$
Entonces

$P(X = 0, Y = 0) = \frac{1}{9}$; pensando de forma análoga los otros casos:

$P(X = 1, Y = 0) = \frac{2}{9}$; $P(X = 2, Y = 0) = \frac{1}{9}$; $P(X = 0, Y = 1) = \frac{2}{9}$, $P(X = 1, Y = 1) = \frac{2}{9}$,

$P(X = 0, Y = 2) = \frac{1}{9}$; $P(X = 1, Y = 2) = P(X = 2, Y = 2) = 0$

Disponemos estas probabilidades en una tabla de la siguiente forma

Y \ X	**0**	**1**	**2**
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

5.2 - Funciones de distribución marginales de una v.a. (*X*,*Y*) discreta

En el ejemplo 1, supongamos que queremos saber cuál es la probabilidad de que el número de artículos producidos por la línea I sea 2, o sea $P(X=2)$
Como el evento $\{X=2\}$ es igual a $\{X=2\}\cap(\{Y=0\}\cup\{Y=1\}\cup\{Y=2\}\cup\{Y=3\})$, y a su vez
$\{X=2\}\cap(\{Y=0\}\cup\{Y=1\}\cup\{Y=2\}\cup\{Y=3\})=$
$=(\{X=2\}\cap\{Y=0\})\cup(\{X=2\}\cap\{Y=1\})\cup(\{X=2\}\cap\{Y=2\})\cup(\{X=2\}\cap\{Y=3\})$
Entonces

$$P(X=2)=$$
$$=P(\{X=2\}\cap\{Y=0\})+P(\{X=2\}\cap\{Y=1\})+P(\{X=2\}\cap\{Y=2\})+P(\{X=2\}\cap\{Y=3\})=$$
$$=P(X=2,Y=0)+P(X=2,Y=1)+P(X=2,Y=2)+P(X=2,Y=3)=\sum_{j=0}^{3}P(X=2,Y=j)$$

Razonando de la misma forma podemos escribir

$$P(X=i)=\sum_{j=0}^{3}P(X=i,Y=j) \qquad i=0,1,...,5$$

Es decir obtenemos la ***función de distribución de probabilidad de X***

Análogamente obtenemos

$$P(Y=j)=\sum_{i=0}^{5}P(X=i,Y=j) \qquad j=0,1,2,3$$

Que es la ***función de distribución de probabilidad de Y***

En general se las denomina ***distribuciones marginales de X e Y***, y su definición sería la siguiente

Sea (*X*,*Y*) discreta y sea $p(x_i,y_j)$ (*i*=1,2,…*n*, *j*=1,2,…,*m*) su función de probabilidad conjunta (Eventualmente *n* y/o *m* pueden ser ∞).

La función de probabilidad marginal de *X* es

$$p(x_i)=P(X=x_i)=\sum_{j=1}^{m}p(x_i,y_j) \qquad (i=1,2,\ldots,n)$$

La función de probabilidad marginal de *Y* es

$$q(y_j)=P(Y=y_j)=\sum_{i=1}^{n}p(x_i,y_j) \qquad (j=1,2,\ldots,m)$$

Observación: Remarcamos que la función de probabilidad marginal de *X*, es decir $p(x_i)$ calculada a partir de $p(x_i,y_j)$ en la forma indicada, coincide con la función de probabilidad de la variable aleatoria unidimensional *X* considerada en forma aislada. Análogamente la función de probabilidad marginal de

Y, es decir $q(y_j)$ calculada a partir de $p(x_i, y_j)$ en la forma indicada, coincide con la función de probabilidad de variable aleatoria unidimensional *Y* considerada en forma aislada.

Ejemplo:
Siguiendo con el ejemplo 1,

$$p(5) = P(X = 5) = p(5,0) + p(5,1) + p(5,2) + p(5,3) = 0.09 + 0.08 + 0.06 + 0.05$$
$$= 0.28$$

$$q(1) = P(Y = 1) = p(0,1) + p(1,1) + p(2,1) + p(3,1) + p(4,1) + p(5,1) = 0.01 + 0.02 + 0.04 + 0.05 + 0.06$$
$$= 0.26$$

Observemos que se verifica la condición de normalización para cada una de las marginales:

$$\sum_{x_i=0}^{5} p(x_i) = 0.03 + 0.08 + 0.16 + 0.21 + 0.24 + 0.28 = 1$$

$$\sum_{y_j=0}^{3} q(y_j) = 0.25 + 0.26 + 0.25 + 0.24 = 1$$

5.3 - Funciones de probabilidades condicionales

Consideremos nuevamente el ejemplo de las dos líneas I y II que producen cierto artículo a pequeña escala. Definimos la v.a. (X, Y) cuya función de probabilidad conjunta $p(x_i, y_j)$ está dada por la tabla anterior que repetimos

Y/X	0	1	2	3	4	5	$q(y_j)$
0	0	0.01	0.03	0.05	0.07	0.09	0.25
1	0.01	0.02	0.04	0.05	0.06	0.08	0.26
2	0.01	0.03	0.05	0.05	0.05	0.06	0.25
3	0.01	0.02	0.04	0.06	0.06	0.05	0.24
$p(x_i)$	0.03	0.08	0.16	0.21	0.24	0.28	1

Supongamos que deseamos conocer la probabilidad de que la línea I produzca tres artículos sabiendo que la línea II ha fabricado dos. Tenemos que calcular una probabilidad condicional. Entonces

$$P(X = 3 | Y = 2) = \frac{P\left(X = 3 \ , \ Y = 2\right)}{P(Y = 2)} = \frac{p(3,2)}{q(2)} = \frac{0.05}{0.25} = 0.2$$

En general definimos la ***función de probabilidad puntual de X condicional a Y*** como sigue:

$p(x_i | y_j) = P(X = x_i | Y = y_j) = \dfrac{p(x_i, y_j)}{q(y_j)}$, es decir como el cociente de la función de probabilidad conjunta de (X, Y) y la función de probabilidad puntual marginal de *Y*.

Análogamente, definimos la ***función de probabilidad puntual de Y condicional a X*** :

$q(y_j|x_i) = P(Y = y_j|X = x_i) = \dfrac{p(x_i, y_j)}{p(x_i)}$, es decir como el cociente de la función de probabilidad puntual conjunta de (X,Y) y la función de probabilidad puntual marginal de *X*.

5.4– Variables aleatorias independientes

Ya se discutió el concepto de independencia entre dos eventos *A* y *B*. Esas mismas ideas podemos trasladarlas en relación a dos variables aleatorias *X* e *Y* que, eventualmente, podemos considerarlas como las componentes de una variable aleatoria bidimensional (X,Y).
De acuerdo con esto, intuitivamente decimos que dos variables, *X* e *Y*, son independientes si el valor que toma una de ellas no influye de ninguna manera sobre el valor que toma la otra. Esto lo establecemos más formalmente:

Sea (X,Y) una variable aleatoria bidimensional discreta. Sea $p(x_i, y_j)$ su *fdp* conjunta y $p(x_i)$ y $q(y_j)$ las correspondientes *fdp* marginales de *X* e *Y*. Decimos que *X* e *Y* son ***variables aleatorias independientes*** si y sólo si

$$p(x_i, y_j) = p(x_i)q(y_j) \quad \forall (x_i, y_j) \in R_{XY}$$

Observación: Notar que para poder afirmar la independencia de *X* e *Y* debe cumplirse la factorización de la *fdp* conjunta como producto de las *fdp* marginales para ***todos*** los pares de valores de la v.a. (X,Y). Por lo tanto, para verificar la independencia es necesario demostrar la validez de la factorización para todos los pares. En cambio, es suficiente encontrar un solo par que no la verifica, para afirmar, de acuerdo con la definición, que las variables *X* e *Y* son no independientes, es decir, que son dependientes. Esto es, para demostrar la dependencia es suficiente con encontrar un solo par que no verifique la factorización señalada.

Vimos que dos sucesos *A* y *B* son independientes si y sólo si $P(A|B) = P(A)$ y $P(B|A) = P(B)$ (donde por supuesto debía ser $P(A) \neq 0$ y $P(B) \neq 0$). En términos de variables aleatorias, esta forma de ver la independencia se manifiesta en la igualdad entre las *fdp* condicionales y las correspondientes *fdp* marginales, como demostramos en este

Teorema
Sea (X,Y) una variable aleatoria bidimensional discreta cuyas *fdp* conjunta, condicionales y marginales son, respectivamente, $p(x_i, y_j)$; $p(x_i|y_j)$, $q(y_j|x_i)$ y $p(x_i)$, $q(y_j)$.
Entonces, *X* e *Y* son variables aleatorias independientes si y sólo si

1) $p(x_i|y_j) = p(x_i) \quad \forall (x_i, y_j) \in R_{XY}$, o

2) $q(y_j|x_i) = q(y_j) \quad \forall (x_i, y_j) \in R_{XY}$, que es equivalente a lo anterior

Dem.)
Demostraremos solamente 1). La equivalencia entre1) y 2) la dejamos como ejercicio.
Para demostrar 1) verificaremos la doble equivalencia entre ésta y la definición de v.a. independientes.
$\Rightarrow$)
Sean X e Y variables aleatorias independientes. Entonces $\forall (x_i, y_j) \in R_{XY}$

$$p(x_i|y_j) = \frac{p(x_i, y_j)}{q(y_j)} = \frac{p(x_i)q(y_j)}{q(y_j)} = p(x_i)$$

Aquí la primera igualdad es la definición de *fdp* condicional y la segunda sale de la definición de independencia al suponer que X e Y son independientes.
$\Leftarrow$)
Supongamos que se verifica 1). Entonces $\forall (x_i, y_j) \in R_{XY}$

$$p(x_i|y_j) = p(x_i) \rightarrow \frac{p(x_i, y_j)}{q(y_j)} = p(x_i) \rightarrow p(x_i, y_j) = p(x_i)q(y_j) \rightarrow X \text{ e } Y \text{ independientes}$$

Aquí, la primera implicación se debe a la definición de *fdp* condicional y la tercera a la definición de v.a. independientes.

Ejemplo:
1- Supongamos que una máquina se usa para un trabajo específico a la mañana y para uno diferente en la tarde. Representemos por X e Y el número de veces que la máquina falla en la mañana y en la tarde respectivamente. Supongamos que la tabla siguiente da la función de probabilidad conjunta $p(x_i, y_j)$ de la variable aleatoria bidimensional discreta (X, Y).

Y/X	0	1	2	$q(y_j)$
0	0.1	0.2	0.2	0.5
1	0.04	0.08	0.08	0.2
2	0.06	0.12	0.12	0.3
$P(x_i)$	0.2	0.4	0.4	1

Deseamos saber si las variables aleatorias X e Y son independientes o dependientes.
Para demostrar que son independientes debemos probar que se verifica $\forall (x_i, y_j) \in R_{XY}$
$p(x_i, y_j) = p(x_i)q(y_j)$ Verificamos directamente que

$$p(0,0) = 0.1 = p(0)q(0) = 0.2 \times 0.5$$
$$p(0,1) = 0.04 = p(0)q(1) = 0.2 \times 0.2$$
$$p(0,2) = 0.06 = p(0)q(2) = 0.2 \times 0.3$$
$$p(1,0) = 0.2 = p(1)q(0) = 0.4 \times 0.5$$
$$p(1,1) = 0.08 = p(1)q(1) = 0.4 \times 0.2$$
$$p(1,2) = 0.12 = p(1)q(2) = 0.4 \times 0.3$$
$$p(2,0) = 0.2 = p(2)q(0) = 0.4 \times 0.5$$
$$p(2,1) = 0.08 = p(2)q(1) = 0.4 \times 0.2$$
$$p(2,2) = 0.12 = p(2)q(2) = 0.4 \times 0.3$$

Luego X e Y son independientes.

Podríamos haber usado las condiciones 1) $p(x_i|y_j) = p(x_i) \quad \forall (x_i, y_j) \in R_{XY}$, o su equivalente
2) $q(y_j|x_i) = q(y_j) \quad \forall (x_i, y_j) \in R_{XY}$. Veamos, como muestra para un solo valor, que se verifica
$p(2|1) = \dfrac{p(2,1)}{q(1)} = \dfrac{0.08}{0.2} = 0.4 = p(2)$. Para demostrar la independencia por este camino habría que demostrar que se cumple la condición para el resto de los pares de valores. Se deja este cálculo como ejercicio optativo.

Observaciones

1- De la definición de las *fdp* marginales, vemos que tanto en el caso discreto como en el continuo, la *fdp* conjunta determina unívocamente las *fdp* marginales. Es decir, si (X,Y) es discreta del conocimiento de la función de probabilidad conjunta $p(x_i, y_j)$ podemos determinar unívocamente las funciones de probabilidad $p(x_i)$ y $q(y_j)$. Sin embargo la inversa no se cumple en general. Es decir del conocimiento de $p(x_i)$ y $q(y_j)$ no se puede, en general, reconstruir $p(x_i, y_j)$ a menos que *X* e *Y* sean variables independientes en cuyo caso es $p(x_i, y_j) = p(x_i)q(y_j)$.

2- El concepto de independencia entre dos variables aleatorias se puede generalizar a *n* variables aleatorias $X_1, X_2, ..., X_n$

5.5 - Función de una variable aleatoria bidimensional

Existen muchas situaciones en las que dado una variable aleatoria bidimensional nos interesa considerar otra variable aleatoria que es función de aquélla. Por ejemplo, supongamos que las variables aleatorias *X* e *Y* denotan la longitud y el ancho, respectivamente, de una pieza, entonces $Z = 2X + 2Y$ es una v.a. que representa el perímetro de la pieza, o la v.a. $W = X.Y$ representa el área de la pieza. Tanto *Z* como *W* ***son variables aleatorias***.

En general, sea S un espacio muestral asociado a un experimento probabilístico ε, sean $X : S \to R$ e $Y : S \to R$ dos variables aleatorias que definen una variable aleatoria bidimensional (X,Y) cuyo recorrido es R_{XY}, y sea una función de dos variables reales $H : R_{XY} \to R$ que a cada elemento (x, y) del recorrido R_{XY} le hace corresponder un número real $z = H(x, y)$, entonces la función compuesta $Z = H(X,Y) : S \to R$ es una variable aleatoria, puesto que a cada elemento $s \in S$ le hace corresponder un número real $z = H[X(s), Y(s)]$. Diremos que la variable aleatoria *Z* es ***función de la variable aleatoria bidimensional (X,Y).***

Algunas variables aleatorias que son función de variables aleatorias bidimensionales son $Z = X + Y$, $Z = X.Y$, $Z = X / Y$, $Z = mín(X,Y)$, $Z = máx(X,Y)$, etc.

Lo anterior se puede generalizar si en lugar de dos variables aleatorias tenemos *n* variables aleatorias $X_1, X_2, ..., X_n$, y $z = H(x_1, x_2, ...x_n)$ es una función de n variables a valores reales.

Ejemplos:
1- Sea $Z \sim B(n,p)$
Podemos escribir a Z como suma de variables aleatorias de la siguiente forma.
Recordar que Z cuenta el número de éxitos en n repeticiones o ensayos del experimento ε
Si definimos

$$X_i = \begin{cases} 1 & si \ \ en \ \ la \ \ i-ésima \ \ repetición \ \ de \ \ \varepsilon \ \ ocurre \ \ éxito \\ 0 & caso \ \ contrario \end{cases} \qquad i = 1,2,...,n$$

Notar que a cada X_i se la puede considerar $B(1,p)$, y además $X_1, X_2,..., X_n$ son independientes
Podemos escribir $Z = X_1 + X_2 + ... + X_n$

2- Sea Z v.a. binomial negativa con parámetros r y p, es decir $Z \sim BN(r,p)$
Si definimos
X_1: "número de repeticiones del experimento requeridos hasta el 1º éxito"
X_2: "número de repeticiones del experimento adicionales requeridos hasta el 2º éxito"
X_3: "número de repeticiones del experimento adicionales requeridos hasta el 3º éxito"
Y en general
X_i: "número de repeticiones del experimento adicionales después del *(i-1)*–ésimo éxito requeridos hasta el *i*-ésimo éxito"
Entonces cada variable tiene ***distribución geométrica con parámetro p*** y $Z = X_1 + X_2 + ... + X_r$
Notar además que $X_1, X_2,..., X_r$ son independientes

<u>Esperanza de una v.a. que es función de una v.a. bidimensional</u>

Sea una variable aleatoria bidimensional (X,Y) cuya *fdp* conjunta es la función de probabilidad conjunta $p(x_i, y_j)$ si es discreta o la función de densidad de probabilidad conjunta $f(x,y)$ si es continua y sea una función real de dos variables $z = H(x,y)$ de manera que podemos definir una variable aleatoria Z que es función de la variable aleatoria bidimensional (X,Y) de la forma $Z = H(X,Y)$. Si la *fdp* de Z es $q(z_i)$, siendo Z discreta, entonces la esperanza matemática de Z es, de acuerdo con la definición general,

$$E(Z) = \sum_{x_i \in R_X} z_i . q(z_i) \qquad (Z \text{ discreta})$$

Nuevamente lo interesante es considerar la posibilidad de evaluar $E(Z)$ sin tener que calcular previamente la *fdp* de Z. El siguiente teorema nos muestra cómo hacerlo.

<u>Teorema</u> Sea (X,Y) una variable aleatoria bidimensional y sea $Z=H(X,Y)$ una variable aleatoria que es función de (X,Y).

Si Z es variable aleatoria discreta que proviene de la variable aleatoria bidimensional discreta (X,Y) cuyo recorrido es R_{XY} y su *fdp* conjunta es $p(x_i, y_j)$, entonces:

$$E(Z) = E[H(X,Y)] = \sum_{(x_i, y_j) \in R_{XY}} H(x_i, y_j) p(x_i, y_j)$$

Dem.) sin demostración

Esperanza de una suma de variables aleatorias

> Sean X e Y dos variables aleatorias arbitrarias. Entonces $E(X+Y)=E(X)+E(Y)$.

Dem.) en el teorema anterior consideramos $H(x,y)=x+y$
Si *(X,Y)* es discreta

$$E(Z)=E[H(X,Y)]=\sum_{(x_i,y_j)\in R_{XY}} H(x_i,y_j)p(x_i,y_j)=\sum_{(x_i,y_j)\in R_{XY}}(x_i+y_j)p(x_i,y_j)=$$

Aplicando la propiedad distributiva y separando en dos sumas

$$E(Z)=\sum_{(x_i,y_j)\in R_{XY}}(x_i+y_j)p(x_i,y_j)=\sum_{(x_i,y_j)\in R_{XY}}x_i\,p(x_i,y_j)+\sum_{(x_i,y_j)\in R_{XY}}y_j\,p(x_i,y_j)=$$

$$=\sum_i\sum_j x_i\,p(x_i,y_j)+\sum_i\sum_j y_j\,p(x_i,y_j)=\sum_i x_i\sum_j p(x_i,y_j)+\sum_j y_j\sum_i p(x_i,y_j)=$$

Pero $\sum_j p(x_i,y_j)=p(x_i)$ y $\sum_i p(x_i,y_j)=q(y_j)$, por lo tanto

$$=\sum_i x_i\,p(x_i)+\sum_j y_j q(y_j)=E(X)+E(Y)$$

Para el caso *(X,Y)* continua <u>sigue siendo válida esta propiedad</u>.

Podemos generalizar la propiedad anterior a un número finito cualquiera de variables aleatorias:

> Sean $X_1, X_2, ..., X_n$ *n* variables aleatorias arbitrarias. Entonces:
>
> $E(X_1+X_2+...+X_n)=E(X_1)+E(X_2)+...+E(X_n)$ o, en notación más concentrada,:
>
> $$E\left(\sum_{i=1}^{n}X_1\right)=\sum_{i=1}^{n}E(X_i)$$

(leeremos: "la esperanza de la suma es la suma de las esperanzas")

Dem.) Se deduce por inducción completa sobre el número *n* de variables aleatorias.

Observación: se deduce que la esperanza verifica la ***propiedad lineal***:

$$E\left(\sum_{i=1}^{n}a_iX_i\right)=\sum_{i=1}^{n}a_iE(X_i).$$

<u>Ejemplos</u>:
1- Vamos a aplicar algunas de las propiedades anteriores para calcular de una manera alternativa la esperanza matemática de una variable aleatoria X distribuida binomialmente.
Sea entonces una v.a. X~B(n,p).

Ya vimos que podemos escribir $X = X_1 + X_2 + ... + X_n$ donde cada X_i se la puede considerar $B(1, p)$, y además $X_1, X_2, ..., X_n$ son independientes
Entonces
$E(X_i) = 1 \times P(X_i = 1) + 0 \times P(X_i = 0) = P(X_i = 1) = p$ para cualquier i
Por lo tanto

$$E(X) = E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n) = \underbrace{p + p + ... + p}_{n \ veces} = np$$

Observación: muchas veces es conveniente descomponer una variable aleatoria como suma de otras más simples para facilitar los cálculos

2- ***Esperanza de una v.a. binomial negativa***
Cuando se trató la v.a. binomial negativa se dijo cuál era su esperanza. Ahora damos una demostración
Sea *X* v.a. binomial negativa con parámetros *r* y *p*, es decir $X \sim BN(r, p)$
Si definimos
X_1: "número de repeticiones del experimento requeridos hasta el 1º éxito"
X_2: "número de repeticiones del experimento adicionales requeridos hasta el 2º éxito"
X_3: "número de repeticiones del experimento adicionales requeridos hasta el 3º éxito"
Y en general
X_i: "número de repeticiones del experimento adicionales después del *(i-1)*–ésimo éxito requeridos hasta el *i*-ésimo éxito"
Entonces cada variable tiene ***distribución geométrica con parámetro p*** y $X = X_1 + X_2 + ... + X_r$
Por lo tanto

$$E(X) = E(X_1 + X_2 + ... + X_r) = E(X_1) + E(X_2) + ... + E(X_r) = \underbrace{\frac{1}{p} + \frac{1}{p} + ... + \frac{1}{p}}_{r \ veces} = r\frac{1}{p} = \frac{r}{p}$$

3- ***Esperanza de una v.a. hipergeométrica***
Si $X \sim H(n,\ M,\ N)$ entonces $E(X) = \dfrac{nM}{N}$
Para facilitar la demostración supongamos que tenemos *N* bolillas en una urna de las cuales *M* son rojas y *N-M* son blancas. Queremos hallar el número esperado de bolillas rojas extraídas
Definimos las variables

$$X_i = \begin{cases} 1 & si \ la \ i-ésima \ bolilla \ roja \ es \ extraída \\ 0 & caso \ contrario \end{cases}$$

Las variables $X_1, X_2, ... X_M$ ***no son independientes***
Se puede escribir $X = X_1 + X_2 + ... + X_M$, además

$$E(X_i) = P(X_i = 1) = \frac{\binom{1}{1}\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N}$$

Por lo tanto

$$E(X)=E(X_1+X_2+...+X_M)=E(X_1)+E(X_2)+...+E(X_{Mr})=\underbrace{\frac{n}{N}+\frac{n}{N}+...+\frac{n}{N}}_{M\ \ veces}=M\frac{n}{N}=\frac{nM}{N}$$

Ejemplo

El espesor X de una cuña de madera (en milímetros) tiene una función de densidad de probabilidad

$$f(x)=\begin{cases}\frac{3}{4}-\frac{3(x-5)^2}{4} & 4\le x\le 6\\ 0 & \text{en otro lado}\end{cases}$$

a) Determine E(X)

b) Si Y denota el espesor de una cuña en pulgadas (1mm = 0.0394 pulgadas), determine E(Y)

c) Si se seleccionan tres cuñas de manera independiente y las apilamos una encima de otra, encuentre la media y la varianza del espesor total.

a) Verifique el lector que

$$E(X)=\int_4^6 x\left(\frac{3}{4}-\frac{3}{4}(x-5)^2\right)dx=5$$

b) $Y=0.0394X$ entonces $E(Y)=E(0.0394X)=0.0394E(X)=0.197$

c) Notar que si X_i: "espesor de cuña i" , $i=1,2,3$ entonces $X=X_1+X_2+X_3$ es el espesor total

Por lo tanto $E(X)=E(X_1+X_2+X_3)=E(X_1)+E(X_2)+E(X_3)=5+5+5=15$

En general ***la esperanza de un producto de variables aleatorias no es igual al producto de las esperanzas***

Si (X,Y) es una variable aleatoria bidimensional tal que X e Y son variables aleatorias ***independientes***, entonces: $E(X.Y)=E(X).E(Y)$

(leeremos:" la esperanza del producto es el producto de las esperanzas").

Dem.) análoga a la demostración de la propiedad anterior.

Para el caso *(X,Y)* continua sigue siendo válida esta propiedad.

Ejemplo:

Supongamos que debido a innumerables causas incontrolables la corriente i y la resistencia r de un circuito varían aleatoriamente de forma tal que pueden considerarse como variables aleatorias I y R independientes. Supongamos que las correspondientes *fdp* son:

$$g(i)=\begin{cases}2i & 0\le i\le 1\\ 0 & demás\ \ valores\end{cases} \qquad h(r)=\begin{cases}\frac{r^2}{9} & 0\le r\le 3\\ 0 & demás\ \ valores\end{cases}$$

Nos interesa considerar el voltaje $v=i.r$ de manera que podemos definir la variable aleatoria $V=I.R$. Hallar el valor esperado o esperanza matemática del voltaje: $E(V)$.

Como I y R son independientes, usando la propiedad anterior

$$E(V)=E(I)E(R)$$

$$E(I)=\int_0^1 i(2i)di = 2\left.\frac{i^3}{3}\right|_0^1 = \frac{2}{3} \qquad E(R)=\int_0^3 r\left(\frac{r^2}{9}\right)dr = \frac{1}{9}\left.\frac{r^4}{4}\right|_0^3 = \frac{1}{9}\times\frac{3^4}{9} = 1$$

$$\therefore \quad E(V)=\frac{2}{3}\times 1 = \frac{2}{3}$$

Varianza de una suma de variables aleatorias

.

$$\boxed{V(X+Y)=V(X)+V(Y)+2\sigma_{XY} \quad \text{con} \quad \sigma_{XY}=E(X.Y)-E(X).E(Y)}$$

Dem.) Escribimos la varianza en su forma alternativa

$V(X+Y)=E\left([X+Y]^2\right)-[E(X+Y)]^2$. Desarrollamos los cuadrados y aplicamos la propiedad lineal de la esperanza:

$$\begin{aligned} V(X+Y) &= E\left(X^2+2.X.Y+Y^2\right)-[E(X)+E(Y)]^2 \\ &= E\left(X^2\right)+2E(X.Y)+E\left(Y^2\right)-\left\{[E(X)]^2+2E(X)E(Y)+[E(Y)]^2\right\} \end{aligned}$$

Agrupando convenientemente:

$$\begin{aligned} V(X+Y) &= \left\{E\left(X^2\right)-[E(X)]^2\right\}+\left\{E\left(Y^2\right)-[E(Y)]^2\right\}+2\{E(X.Y)-E(X)E(Y)\} \\ &= V(X)+V(Y)+2\{E(X.Y)-E(X)E(Y)\} \end{aligned}$$

, es decir

$$V(X+Y)=V(X)+V(Y)+2\sigma_{XY}$$

$\boxed{\sigma_{XY}=E(X.Y)-E(X).E(Y)}$ se la llama la ***covarianza de X e Y.***

Observaciones:

1- Teniendo presente la definición de la desviación estándar de una v.a. *X:* $\sigma_X=\sqrt{V(X)}$, vemos que a la propiedad anterior la podemos escribir:

$$V(X+Y)=\sigma_X^2+\sigma_Y^2+2\sigma_{XY}$$

2- Análogamente se prueba que $V(X-Y)=\sigma_X^2+\sigma_Y^2-2\sigma_{XY}$

3- X e Y son independientes, entonces $V(X+Y)=V(X-Y)=V(X)+V(Y)$

Esto es porque si las variables aleatorias X e Y son independientes, entonces $E(X.Y)=E(X).E(Y)$.

Por lo tanto la covarianza vale cero : $\sigma_{XY}=E(X.Y)-E(X).E(Y)=0$.

4- Podemos generalizar, usando el principio de inducción completa, al caso de n variables aleatorias independientes:
Si $X_1, X_2, ..., X_n$ son n variables aleatorias independientes entonces:

$V(X_1 + X_2 + ... + X_n) = V(X_1) + V(X_2) + ... + V(X_n)$ o, en forma más compacta, $V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i)$.

5- Vemos que la esperanza de la suma de dos variables aleatorias X e Y es igual a la suma de las esperanzas $E(X+Y) = E(X) + E(Y)$ cualesquiera sean X e Y. En cambio la varianza de la suma de las variables aleatorias X e Y es, en general, igual a la suma de las varianzas, $V(X+Y) = V(X) + V(Y)$, sólo si X e Y son variables independientes.

Ejemplos:
1- Podemos ejemplificar la aplicación de las propiedades de la varianza, calculando nuevamente la varianza de una v.a. X distribuida binomialmente con parámetros n y p.
Sea entonces una v.a. X~B(n,p). Vimos que se puede escribir:
$X = X_1 + X_2 + ... + X_n$, donde las n variables aleatorias son independientes entre sí y tienen todas la misma distribución:
$X_i \sim B(1, p) \;\; \forall i = 1,2,...,n$
Entonces, tratándose de n variables aleatorias independientes
$V(X) = V(X_1) + V(X_2) + ... + V(X_n)$ todas la varianzas son iguales y podemos escribir la suma como n veces una cualquiera de ellas:
$V(X) = nV(X_i)$. Pero
$V(X_i) = E(X_i^2) - [E(X_i)]^2$.
Ya vimos que $E(X_i) = 1.p + 0(1-p) = 0$
Además es: $E(X_i^2) = 1^2.p + 0^2(1-p) = p$
Entonces: $V(X_i) = E(X_i^2) - [E(X_i)]^2 = p - p^2 = p(1-p)$.
Luego:
$V(X) = nV(X_i) = np(1-p)$
que es el resultado que habíamos obtenido a partir de la definición y llevando las sumas involucradas a la forma del desarrollo de un binomio de Newton.

2- ***Varianza de una v.a. binomial negativa***
Ya vimos que podemos escribir $X = X_1 + X_2 + ... + X_r$, donde cada variable X_i tiene ***distribución geométrica con parámetro p***
Por lo tanto

$$V(X) = V(X_1) + V(X_2) + ... + V(X_r) = r\frac{1-p}{p^2}$$

5.6 - Covarianza

Sean X e Y dos variables aleatorias. La ***covarianza de X e Y*** se define:

$$\boxed{Cov(X,Y) = E\{[X - E(X)].[Y - E(Y)]\} = E(X.Y) - E(X).E(Y)}$$

Notación: la notación usual para la covarianza de X e Y es σ_{XY} o $Cov(X,Y)$

La última igualdad surge de desarrollar el producto y aplicar las propiedades de la esperanza:
$E\{[X-E(X)][Y-E(Y)]\}=E\{X.Y-X.E(Y)-E(X).Y+E(X)E(Y)\}$

Teniendo presente que $E(X)$ y $E(Y)$ son constantes:
$E\{[X-E(X)][Y-E(Y)]\}=E(X.Y)-E(X).E(Y)-E(X)E(Y)+E(X)E(Y)=E(X.Y)-E(X).E(Y)$.

Si X e Y son variables aleatorias independientes, entonces $Cov(X,Y)=0$.

Dem.)
Según vimos, si X e Y son variables aleatorias independientes, entonces $E(X.Y)=E(X).E(Y)$, de donde se sigue la propiedad.

Propiedades de la covarianza

Las siguientes propiedades son útiles y su verificación se deja como ejercicio
1- $Cov(a+bX,c+dY)=bdCov(X,Y)$
2- $Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)$
3- $Cov\left(\sum_{i=1}^{n}X_i,\sum_{j=1}^{m}Y_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{m}Cov(X_i,Y_j)$
4- $Cov(X,X)=V(X)$

Ejemplos:
1)***Varianza de una v.a. hipergeométrica***
Si $X\sim H(n,\ M,\ N)$ entonces $V(X)=n\dfrac{M}{N}\left(\dfrac{N-M}{N}\right)\left(\dfrac{N-n}{N-1}\right)$
Para facilitar la demostración supongamos que tenemos N bolillas en una urna de las cuales M son rojas y N-M son blancas. Queremos hallar la varianza del número de bolillas blancas extraídas
Como antes definimos las variables

$$X_i=\begin{cases}1 \quad si \quad la \quad i-ésima \quad bolilla \quad roja \quad es \quad extraída \\ \qquad 0 \quad caso \quad contrario\end{cases}$$

Las variables $X_1,X_2,...X_M$ ***no son independientes***
Se puede escribir $X=X_1+X_2+...+X_M$, además

$$E(X_i)=P(X_i=1)=\frac{\binom{1}{1}\binom{N-1}{n-1}}{\binom{N}{n}}=\frac{n}{N} \qquad \text{y}$$

$$V(X_i)=E(X_i^2)-(E(X_i))^2=\frac{n}{N}-\left(\frac{n}{N}\right)^2=\frac{n}{N}\left(1-\frac{n}{N}\right)$$

Por lo tanto $V(X) = V(X_1 + X_2 + ... + X_M) = \sum_{i=1}^{M} V(X_i) + 2 \sum_{i1 \le < j \le M} Cov(X_i, X_j)$

Por otro lado $Cov(X_i; X_j) = E(X_i X_j) - E(X_i)E(Y_j)$

Y $E(X_i X_j) = \dfrac{n(n-1)}{N(N-1)}$, entonces $Cov(X_i; X_j) = \dfrac{n(n-1)}{N(N-1)} - \left(\dfrac{n}{N}\right)^2$

Aplicando algunos pasos algebraicos se llega a $Cov(X_i; X_j) = \dfrac{n(n-1)}{N(N-1)} - \left(\dfrac{n}{N}\right)^2 = -\dfrac{n}{N}\left(\dfrac{1}{N-1}\right)\left(1 - \dfrac{n}{N}\right)$

Reemplazando

$$V(X) = \sum_{i=1}^{M} V(X_i) + 2 \sum_{i1 \le < j \le M} Cov(X_i, X_j) = M \frac{n}{N}\left(1 - \frac{n}{N}\right) + 2\left[\binom{M}{2}\left(-\frac{1}{N-1}\right)\frac{n}{N}\left(1 - \frac{n}{N}\right)\right]$$

Nuevamente, luego de algunos cálculos algebraicos se llega a

$$V(X) = n\frac{M}{N}\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)$$

2) De una caja con frutas que contiene 3 naranjas, 2 manzanas y 3 plátanos se selecciona una muestra de 4 frutas.
Sean las variables aleatorias
X: " nº de naranjas extraídas"
Y: "nº de manzanas extraídas"
Notar que la *f.d.p.* conjunta de *(X,Y)* es

$$P(X = x, Y = y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{4-x-y}}{\binom{8}{4}} \qquad x = 0,\ 1,\ 2,\ 3\ ;\ y = 0,\ 1,\ 2\ \ ;\ 1 \le x + y \le 4$$

Es un ejemplo de ***v.a. hipergeométrica bidimensional***.
También se podría haber presentado la *f.d.p.* conjunta en una tabla, donde también figuran las distribuciones marginales de *X* e *Y*.

X/Y	**0**	**1**	**2**	
0	0	2/70	3/70	5/70
1	3/70	18/70	9/70	30/70
2	9/70	18/70	3/70	30/70
3	3/70	2/70	0	5/70
	15/70	40/70	15/70	

a) ¿Cuales son *E(X)* , *V(X), E(Y)* y *V(Y)?*

$$E(X) = 0 \times \frac{0}{70} + 1 \times \frac{3}{70} + 2 \times \frac{9}{70} + 3 \times \frac{3}{70} =$$

$$= \frac{105}{70} = \frac{3}{2}$$

$$E(Y) = 0 \times \frac{15}{70} + 1 \times \frac{40}{70} + 2 \times \frac{15}{70} = 1$$

Verifique el lector que $V(X) = \dfrac{15}{28}$ y $V(Y) = \dfrac{3}{7}$

b) ¿Son *X* e *Y* independientes?

$P(X=0,Y=0)=0$ pero $P(X=0)P(Y=0)=\frac{5}{70}\frac{15}{70}\neq 0$

Por lo tanto X e Y son dependientes, lo que implica que $Cov(X,Y)\neq 0$

c) ¿Cuál es la *Cov(X,Y)?*

$$Cov(X,Y)=E(XY)-E(X)E(Y)$$

$$\begin{aligned}E(XY)&=0\times 0\times 0+0\times 1\times\frac{2}{70}+0\times 2\times\frac{3}{70}+\\&+1\times 0\times\frac{3}{70}+1\times 1\times\frac{18}{70}+1\times 2\times\frac{9}{70}+\\&+2\times 0\times\frac{9}{70}+2\times 1\times\frac{18}{70}+2\times 2\times\frac{3}{70}+\\&+3\times 0\times\frac{3}{70}+3\times 1\times\frac{2}{70}+3\times 2\times 0=\frac{90}{70}\end{aligned}$$

Entonces $Cov(X,Y)=E(XY)-E(X)E(Y)=\frac{9}{7}-\frac{3}{2}\times 1=-\frac{3}{14}$

d) Z = X+Y simboliza el total de naranjas y manzanas extraídas

¿Cuál es la *E(Z)* y *V(Z)* ?

$$E(Z)=E(X+Y)=E(X)+E(Y)=\frac{3}{2}+1=2.5$$

$$V(Z)=V(X)+V(Y)+2Cov(X,Y)=\frac{15}{18}+\frac{3}{7}+2\times\left(-\frac{3}{14}\right)=\frac{15}{28}$$

e) Supongamos que cada naranja cuesta 2$ y cada manzana cuesta 1.5$ entonces

$W=2X+1.5Y$ es el costo del total de frutas extraídas. Hallar *E(W)* y *V(W)*

$$E(W)=E(2X+1.5Y)=2E(X)+1.5E(Y)=4.5\$$$

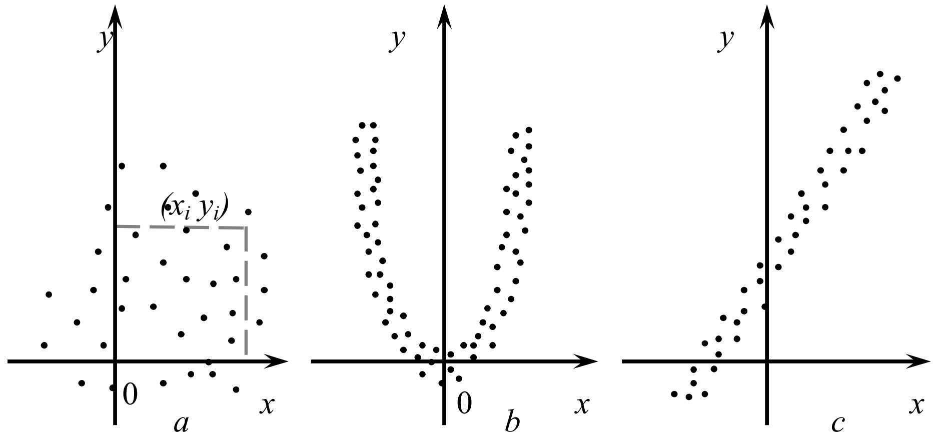
$$V(W)=V(2X+1.5Y)=2^2V(X)+1.5^2V(Y)+2\times 2\times 1.5Cov(X,Y)=\frac{51}{28}$$

5.7 - Coeficiente de correlación lineal.

En realidad más que la covarianza aquí nos interesa considerar una cantidad relacionada con σ_{XY} y que según veremos nos dará información sobre el grado de asociación que existe entre X e Y. Más concretamente nos contará si existe algún grado de relación lineal entre X e Y. Esa cantidad es el coeficiente de correlación lineal.

En el mismo sentido en que podemos tener una idea aproximada sobre la probabilidad de un suceso *A* si repetimos el experimento y consideramos las ocurrencias de *A* en las *n* repeticiones,

así podemos tener también una primera idea sobre la existencia de una relación funcional, específicamente una relación lineal, entre X e Y si consideramos un ***diagrama de dispersión***. Consiste en dibujar pares de valores (x_i,y_j) medidos de la variable aleatoria (X,Y) en un sistema de coordenadas. En la figura mostramos diversas situaciones posibles.



De la figura *a* se deduciría que entre X e Y no hay ningún tipo de relación funcional. La figura *b* sugiere la posibilidad de que exista una relación funcional que corresponde a una parábola. La figura *c*, por su parte, sugiere una relación lineal entre X e Y. Este último es el comportamiento que nos interesa caracterizar. Con ese fin definimos el coeficiente de correlación lineal como sigue:

> Sea (X,Y) una variable aleatoria bidimensional. Definimos el ***coeficiente de correlación lineal entre X e Y*** como $\rho_{XY} = \dfrac{Cov(X,Y)}{\sigma_X \sigma_Y}$

En consecuencia:

$$\rho_{XY} = \frac{E\{[X - E(X)].[Y - E(Y)]\}}{\sqrt{V(X).V(Y)}} = \frac{E(X.Y) - E(X).E(Y)}{\sqrt{V(X).V(Y)}}.$$

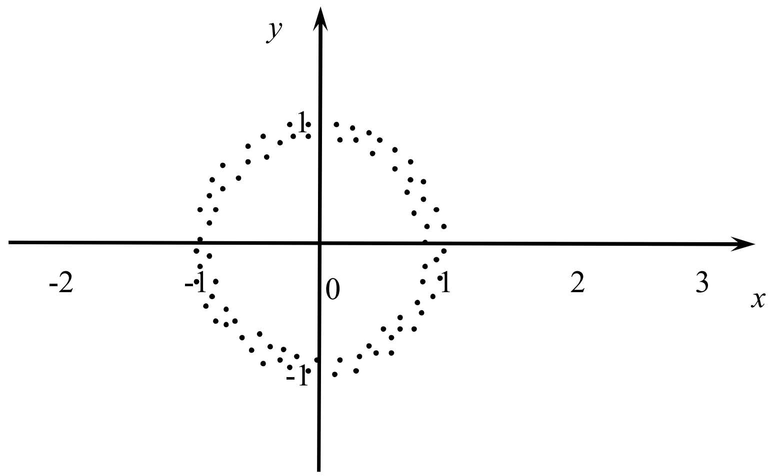
Daremos una serie de propiedades de ρ_{XY} que nos permitirán establecer más concretamente su significado.

<u>Propiedad 1</u>

> Si X e Y son variables aleatorias independientes entonces $\rho_{XY} = 0$.

Dem.) inmediata a partir del hecho que si X e Y son independientes entonces $E(XY) = E(X)E(Y)$

<u>Observación:</u> La inversa no es necesariamente cierta. Puede ser que $\rho_{XY} = 0$ y sin embargo X e Y no sean variables aleatorias independientes. En efecto si tenemos una v.a. bidimensional (X,Y) que da lugar a un diagrama de dispersión como el que se muestra en la figura, veremos que correspondería a un coeficiente de correlación lineal $\rho_{XY} = 0$ y sin embargo la figura sugiere que entre X e Y existe la relación funcional $X^2 + Y^2 = 1$, es decir X e Y son v.a. dependientes. En realidad, como veremos, ρ_{XY} es una medida de la existencia de una relación lineal entre X e Y y una circunferencia se aleja mucho de una línea recta.



Propiedad 2 :

$$\boxed{-1 \leq \rho_{XY} \leq 1}$$

Dem.)

Si consideramos la v.a. $\dfrac{X}{\sigma_X}+\dfrac{Y}{\sigma_Y}$ entonces

$$0 \leq V\left(\frac{X}{\sigma_X}+\frac{Y}{\sigma_Y}\right)=\frac{V(X)}{\sigma_X{}^2}+\frac{V(Y)}{\sigma_Y{}^2}+\frac{2Cov(X,Y)}{\sigma_X\sigma_Y}=2(1+\rho_{XY})$$

Implicando que $-1 \leq \rho_{XY}$

Por otro lado: $0 \leq V\left(\dfrac{X}{\sigma_X}-\dfrac{Y}{\sigma_Y}\right)=\dfrac{V(X)}{\sigma_X{}^2}+\dfrac{V(Y)}{\sigma_Y{}^2}-\dfrac{2Cov(X,Y)}{\sigma_X\sigma_Y}=2(1-\rho_{XY})$

Implicando que $\rho_{XY} \leq 1$

$$\therefore \quad -1 \leq \rho_{XY} \leq 1$$

Propiedad 3 :

Si $\rho_{XY}^2=1$, entonces con probabilidad 1 es $Y = A.X + B$ donde A y B son constantes.

Dem.) Si $\rho_{XY}^2=1$ entonces $\rho_{XY}=1$ o $\rho_{XY}=-1$

Si $\rho_{XY}=-1$ entonces de la demostración anterior se deduce que

$V\left(\dfrac{X}{\sigma_X}+\dfrac{Y}{\sigma_Y}\right)=2(1+\rho_{XY})=0$, lo que implica que la v.a. $Z=\dfrac{X}{\sigma_X}+\dfrac{Y}{\sigma_Y}$ tiene varianza cero. Según la interpretación de varianza podemos deducir (en forma intuitiva) que la v.a. ***no tiene dispersión con respecto a su esperanza***, es decir la v.a. ***Z es una constante con probabilidad 1***

Por lo tanto esto implica que $Y = A.X + B$ con $A=-\dfrac{\sigma_Y}{\sigma_X}<0$

Análogamente $\rho_{XY}=1$ implica que $Y = A.X + B$ con $A=\dfrac{\sigma_Y}{\sigma_X}>0$

Propiedad 4 :

Si *X* e *Y* son dos variables aleatorias tales que $Y = A.X + B$, donde *A* y *B* son constantes, entonces $\rho_{XY}^2=1$. Si $A>0$ es $\rho_{XY}=1$ y si $A<0$ es $\rho_{XY}=-1$.

Dem.) se deja como ejercicio

Observación: Claramente las propiedades anteriores establecen que el coeficiente de correlación lineal es una medida del grado de linealidad entre *X* e *Y*.

Ejemplo

En el ejemplo anterior

$$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}}=\frac{-\frac{3}{14}}{\sqrt{\frac{15}{28}\,\frac{3}{7}}}=-\frac{\sqrt{5}}{5}=-0.44721$$

6- SUMA DE VARIABLES ALEATORIAS Y TEOREMA CENTRAL DEL LÍMITE

6.1 – Suma de variables aleatorias independientes

Cuando se estudiaron las variables aleatorias bidimensionales se habló de una ***función de variable aleatoria bidimensional.*** En particular se nombró la suma de n variables aleatorias**,** pero no se dijo nada sobre la ***distribución*** de esa v.a. suma.
Es a menudo importante saber cuál es la distribución de una suma de variables aleatorias independientes.
Consideramos algunos ejemplos en el caso discreto

1- ***Suma de variables aleatorias independientes con distribución Poisson***

$$X \sim P(\lambda_1)\ ;\ Y \sim P(\lambda_2)\ ;\ X \text{ y } Y \text{ independientes} \quad \Rightarrow \quad X+Y \sim P(\lambda_1+\lambda_2)$$

Dem.)
Consideramos el evento $\{X+Y=n\}$ como unión de eventos excluyentes $\{X=k, Y=n-k\}$ $\quad 0 \le k \le n$, entonces

$$P(X+Y=n)=\sum_{k=0}^{n} P(X=k,Y=n-k)=\sum_{k=0}^{n} P(X=k)P(Y=n-k)=\sum_{k=0}^{n} e^{-\lambda_1}\frac{\lambda_1^{\,k}}{k!}e^{-\lambda_2}\frac{\lambda_2^{\,n-k}}{(n-k)!}=$$

X e Y independientes

$$=e^{-(\lambda_1+\lambda_2)}\sum_{k=0}^{n}\frac{\lambda_1^{\,k}}{k!}\frac{\lambda_2^{\,n-k}}{(n-k)!}=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\lambda_1^{\,k}\lambda_2^{\,n-k}=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1+\lambda_2)^n$$

Binomio de Newton

O sea $X+Y$ tiene distribución Poisson con parámetro $\lambda_1+\lambda_2$

2- ***Suma de variables aleatorias binomiales independientes***

$$X \sim B(n_1,p)\ ;\ Y \sim B(n_2,p)\ ;\ X \text{ y } Y \text{ independientes} \quad \Rightarrow \quad X+Y \sim B(n_1+n_2,p)$$

Dem.)
Nuevamente consideramos el evento $\{X+Y=k\}$ como unión de eventos excluyentes $\{X=i, Y=k-i\}$ $\quad 0 \le i \le n_1$, entonces

$$P(X+Y=k)=\sum_{i=0}^{n_1} P(X=i,Y=k-i)=\sum_{i=0}^{n_1} P(X=i)P(Y=k-i)=\sum_{k=0}^{n_1}\binom{n_1}{i}p^i(1-p)^{n_1-i}\binom{n_2}{k-i}p^{k-i}(1-p)^{n_2-k+i}=$$

X e Y independientes

$$=p^k(1-p)^{n_1+n_2-k}\sum_{i=0}^{n_1}\binom{n_1}{i}\binom{n_2}{k-i}$$

En la expresión anterior si $j>r$ entonces $\binom{r}{j}=0$

Por último usamos la siguiente identidad combinatoria $\sum_{i=0}^{n_1}\binom{n_1}{i}\binom{n_2}{k-i}=\binom{n_1+n_2}{k}$

Y entonces

$$P(X+Y=k)=\binom{n_1+n_2}{k}p^k(1-p)^{n_1+n_2-k}$$

O sea *X+Y* tiene distribución binomial con parámetros n_1+n_2 y p

<u>Observación</u>: en los dos casos anteriores se puede generalizar el resultado a *n* variables aleatorias independientes, usando el principio de inducción completa, es decir

1- Si $X_1, X_2, \dots, X_n$ son *n* variables aleatorias independientes donde $X_i \sim P(\lambda_i)$ para todo $i=1,2,\dots,n$ entonces $\sum_{i=0}^{n} X_i \sim P(\sum_{i=0}^{n}\lambda_i)$

2- Si $X_1, X_2, \dots, X_n$ son *n* variables aleatorias independientes donde $X_i \sim B(n_i, p)$ para todo $i=1,2,\dots,n$ entonces $\sum_{i=0}^{n} X_i \sim B(\sum_{i=0}^{n} n_i, p)$

<u>Suma de variables aleatorias normales independientes</u>

Si *X* e *Y* son dos variables aleatorias continuas independientes con densidades *g(x)* y *h(y)* respectivamente se puede probar (no lo demostraremos aquí) que la v.a. $Z=X+Y$ tiene densidad dada por $f_{X+Y}(z)=\int_{-\infty}^{\infty} g(z-y)h(y)dy$

Usando esto se puede demostrar el siguiente importante resultado:

Si *X* e *Y* son variables aleatorias independientes donde $X \sim N(\mu_1, \sigma_1{}^2)$ y $Y \sim N(\mu_2, \sigma_2{}^2)$ entonces $X+Y \sim N(\mu_1+\mu_2, \sigma_1{}^2+\sigma_2{}^2)$

Por inducción completa se puede generalizar este resultado a *n* variables:

Si $X_1, X_2, \dots, X_n$ son *n* variables aleatorias independientes donde $X_i \sim N(\mu_i, \sigma_i^2)$ para todo $i=1,2,\dots,n$ entonces $\sum_{i=0}^{n} X_i \sim N(\sum_{i=0}^{n}\mu_i, \sum_{i=1}^{n}\sigma_i^2)$

De lo anterior y del hecho que $X \sim N(\mu, \sigma^2) \Rightarrow aX+b \sim N(a\mu+b, a^2\sigma^2)$ tenemos:

Si $X_1, X_2, \dots, X_n$ son *n* variables aleatorias independientes donde $X_i \sim N(\mu_i, \sigma_i^2)$ para todo $i=1,2,\dots,n$ entonces $\sum_{i=0}^{n} a_i X_i \sim N(\sum_{i=0}^{n} a_i\mu_i, \sum_{i=1}^{n} a_i^2\sigma_i^2)$ donde $a_1, a_2, \dots, a_n$ son números reales

Se dice que $\sum_{i=0}^{n} a_i X_i$ es una ***combinación lineal de variables aleatorias***.

Ejemplos:

1- La envoltura de plástico para un disco magnético está formada por dos hojas. El espesor de cada una tiene una distribución normal con media 1.5 milímetros y desviación estándar de 0.1 milímetros. Las hojas son independientes.

a) Determine la media y la desviación estándar del espesor total de las dos hojas.

b) ¿Cuál es la probabilidad de que el espesor total sea mayor que 3.3 milímetros?

Solución: Sean las variables aleatorias

X: "espesor de la hoja 1" e *Y*: "espesor de la hoja 2"

Entonces $X \sim N(1.5, 0.1^2)$; $Y \sim N(1.5, 0.1^2)$ y X e Y independientes

a) Si definimos la v.a. *Z*: "espesor total de las dos hojas" , entonces $Z = X + Y$

Por lo tanto $Z \sim N(1.5 + 1.5, \quad 0.1^2 + 0.1^2)$ es decir $Z \sim N(3, \quad 0.02)$

En consecuencia $E(Z) = 3$, $\sigma_Z = \sqrt{V(Z)} = \sqrt{0.02}$

b) Se pide calcular $P(Z > 3.3)$

$$P(Z > 3.3) = P\left(\frac{Z-3}{\sqrt{0.02}} > \frac{3.3-3}{\sqrt{0.02}}\right) = 1 - \Phi\left(\frac{3.3-3}{\sqrt{0.02}}\right) = 1 - \Phi(2.12132) = 1 - 0.983 = 0.017$$

2-Tengo tres mensajes que atender en el edificio administrativo. Sea X_i : " el tiempo que toma el *i*-ésimo mensaje" ($i = 1, 2, 3$), y sea X_4 : " el tiempo total que utilizo para caminar hacia y desde el edificio y entre cada mensaje". Suponga que las X_i son independientes, normalmente distribuidas, con las siguientes medias y desviaciones estándar:

$$\mu_1 = 15\,\text{min}, \quad \sigma_1 = 4, \quad \mu_2 = 5, \quad \sigma_2 = 1, \quad \mu_3 = 8, \quad \sigma_3 = 2, \quad \mu_4 = 12, \quad \sigma_4 = 3$$

Pienso salir de mi oficina precisamente a las 10.00 a.m. y deseo pegar una nota en mi puerta que dice "regreso a las *t* a.m." ¿A qué hora *t* debo escribir si deseo que la probabilidad de mi llegada después de *t* sea 0.01?

Solución: Definimos la v.a. *Z*: "tiempo transcurrido desde que salgo de mi oficina hasta que regreso", entonces $T = X_1 + X_2 + X_3 + X_4$

Por lo tanto $T \sim N\left(\sum_{i=1}^{4} \mu_i, \quad \sum_{i=1}^{4} \sigma_i^2\right)$, y se pide hallar *t* tal que $P(T > t) = 0.01$

$$\sum_{i=1}^{4} \mu_i = 15 + 5 + 8 + 12 = 50 \quad \text{y} \quad \sum_{i=1}^{4} \sigma_i^2 = 4^2 + 1^2 + 2^2 + 3^2 = 30$$

Entonces $P(T > t) = 1 - \Phi\left(\frac{t-50}{\sqrt{30}}\right) = 0.01$, es decir $\Phi\left(\frac{t-50}{\sqrt{30}}\right) = 0.99$

Buscando en la tabla de la normal $\frac{t-50}{\sqrt{30}} = 2.33 \quad \Rightarrow \quad t = 2.33 \times \sqrt{30} + 50 = 62.7619$

3- El ancho del marco de una puerta tiene una distribución normal con media 24 pulgadas y desviación estándar de 1/8 de pulgada. El ancho de la puerta tiene una distribución normal con media 23.875 de pulgadas y desviación estándar de 1/16 de pulgadas. Suponer independencia.

a) Determine la distribución, la media y la desviación estándar de la diferencia entre el ancho del marco y de la puerta.

b) ¿Cuál es la probabilidad de que la diferencia entre el ancho del marco y de la puerta sea mayor que ¼ de pulgada?.
c) ¿Cuál es la probabilidad de que la puerta no quepa en el marco?.

Solución: Sean las variables aleatorias
X: "ancho del marco de la puerta en pulgadas"
Y: "ancho de la puerta en pulgadas"
Entonces $X \sim N(24,\ (1/8)^2)$, $Y \sim N(23.875,\ (1/16)^2)$, *X* e *Y* independientes

a) Se pide la distribución de *X-Y* , $E(X-Y)$, $\sigma_{X-Y} = \sqrt{V(X-Y)}$

$$E(X-Y) = E(X) - E(Y) = 24 - 23.875 = 0.125$$

$$V(X-Y) = V(X) + V(Y) = \left(\frac{1}{8}\right)^2 + \left(\frac{1}{16}\right)^2 = \frac{5}{256} \quad \therefore \ \sigma_{X-Y} = \frac{\sqrt{5}}{16}$$

Por lo tanto $X - Y \sim N\left(0.125,\ \left(\frac{\sqrt{5}}{16}\right)^2\right)$

b) Se pide la probabilidad $P(X - Y > 1/4)$

$$P(X-Y > 1/4) = 1 - \Phi\left(\frac{0.25 - 0.125}{\frac{\sqrt{5}}{16}}\right) = 1 - \Phi\left(\frac{2\sqrt{5}}{5}\right) = 1 - \Phi(0.8944) = 1 - 0.8133 = 0.1867$$

c) Si la puerta no entra en el marco entonces se da el evento $\{X < Y\}$ o equivalentemente $\{X - Y < 0\}$, por lo tanto

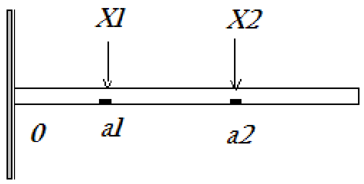
$$P(X-Y<0) = \Phi\left(\frac{0 - 0.125}{\frac{\sqrt{5}}{16}}\right) = \Phi\left(-\frac{2\sqrt{5}}{5}\right) = 1 - \Phi\left(\frac{2\sqrt{5}}{5}\right) = 0.1867$$

4- Supongamos que las variables aleatorias *X* e *Y* denotan la longitud y el ancho en cm, respectivamente, de una pieza.
Supongamos además que *X* e *Y* son independientes y que $X \sim N(2\ ,0.1^2)$, $Y \sim N(5\ ,0.2^2)$.
Entonces $Z = 2X + 2Y$ es una v.a. que representa el perímetro de la pieza.
Calcular la probabilidad de que el perímetro sea mayor que 14.5 cm.

Solución: tenemos que $Z \sim N\left(2\times 2 + 2\times 5,\ 2^2 \times 0.1^2 + 2^2 \times 0.2^2\right)$, o sea $Z \sim N(14,\ 0.2)$
La probabilidad pedida es $P(Z > 14.5)$, entonces

$$P(Z > 14.5) = 1 - \Phi\left(\frac{14.5 - 14}{\sqrt{0.2}}\right) = 1 - \Phi\left(\frac{\sqrt{5}}{2}\right) = 1 - \Phi(1.1180) = 1 - 0.8810 = 0.119$$

5- Si se aplican dos cargas aleatorias X_1 y X_2 a una viga voladiza como se muestra en la figura siguiente, el momento de flexión en 0 debido a las cargas es $a_1 X_1 + a_2 X_2$.



a) Suponga que X_1 y X_2 son v.a. independientes con medias 2 y 4 KLbs respectivamente, y desviaciones estándar 0.5 y 1.0 KLbs, respectivamente.

Si $a_1 = 5$ pies y $a_2 = 10$ pies, ¿cuál es el momento de flexión esperado y cuál es la desviación estándar del momento de flexión?

b) Si X_1 y X_2 están normalmente distribuidas, ¿cuál es la probabilidad de que el momento de flexión supere 75 KLbs?

Solución: Sea la v.a. Z: "momento de flexión en 0", entonces $Z = 5X_1 + 10X_2$

Por lo tanto

a) $E(Z) = 5E(X_1) + 10E(X_2) = 5\times 2 + 10\times 4 = 50$

$$V(Z) = 5^2 \times 0.5^2 + 10^2 \times 1^2 = 25\times 0.25 + 10\times 1 = \frac{65}{4} \quad \therefore \sigma_Z = \sqrt{\frac{65}{4}}$$

b) Si X_1 y X_2 están normalmente distribuidas, entonces $Z \sim N\left(50, \dfrac{65}{4}\right)$

Por lo tanto

$$P(Z > 75) = 1 - \Phi\left(\frac{75-50}{\sqrt{\frac{65}{4}}}\right) = 1 - \Phi\left(\frac{10\sqrt{65}}{13}\right) = 1 - \Phi(6.20) \approx 1 - 1 = 0$$

Promedio de variables aleatorias normales independientes

Si $X_1, X_2, \ldots, X_n$ son n variables aleatorias independientes donde $X_i \sim N(\mu, \sigma^2)$ para todo $i = 1,2,\ldots,n$ entonces la v.a. $\overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ tiene distribución normal con media μ y varianza $\dfrac{\sigma^2}{n}$

Dem.) Notar que $\overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ es un caso particular de combinación lineal de variables aleatorias donde $a_i = \dfrac{1}{n}$ para todo $i = 1,2,\ldots,n$

Además en este caso $\mu_i = \mu$ y $\sigma_i^2 = \sigma^2$ para todo $i = 1,2,\ldots,n$

Por lo tanto, $\overline{X}$ tiene distribución normal con esperanza $\sum_{i=1}^{n} \dfrac{1}{n}\mu_i = \sum_{i=1}^{n} \dfrac{1}{n}\mu = \dfrac{1}{n} n\mu = \mu$ y varianza

$$\sum_{i=1}^{n}\left(\frac{1}{n}\right)^2 \sigma_i^2 = \sum_{i=1}^{n}\left(\frac{1}{n}\right)^2 \sigma^2 = \left(\frac{1}{n}\right)^2 n\sigma^2 = \frac{\sigma^2}{n}$$

Es decir, $\overline{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$

Observación: a $\overline{X}$ se lo llama ***promedio muestral*** o ***media muestral***

Ejemplos:
1) El diámetro interno de un anillo de pistón seleccionado al azar es una v.a. con distribución normal con media 12 cm y desviación estándar de 0.04 cm.
a) Si $\overline{X}$ es el diámetro promedio en una muestra de $n=16$ anillos, calcule $P(11.99 \le \overline{X} \le 12.01)$
b) ¿Qué tan probable es que el diámetro promedio exceda de 12.01 cuando $n=25$?

Solución:
a) Sean las variables aleatorias X_i : "diámetro del anillo i" $i=1,2,\dots,16$

Entonces $X_i \sim N(12, 0.04^2)$ para cada i.

Por lo tanto $\overline{X} \sim N\left(12, \dfrac{0.04^2}{16}\right)$. Entonces

$$P(11.99 \le \overline{X} \le 12.01) = P\left(\frac{11.99-12}{\sqrt{\frac{0.04^2}{16}}} \le \frac{\overline{X}-12}{\sqrt{\frac{0.04^2}{16}}} \le \frac{12.01-12}{\sqrt{\frac{0.04^2}{16}}}\right) =$$

$$= \phi\left(\frac{12.01-12}{\sqrt{\frac{0.04^2}{16}}}\right) - \phi\left(\frac{11.99-12}{\sqrt{\frac{0.04^2}{16}}}\right) = \phi(1) - \phi(-1) = 2\phi(1) - 1 =$$

$$= 2 \times 0.8413 - 1 = 0.6826$$

b) En este caso $\overline{X} \sim N\left(12, \dfrac{0.04^2}{25}\right)$, entonces

$$P(\overline{X} > 12.01) = 1 - \phi\left(\frac{12.01-12}{\sqrt{\frac{0.04^2}{25}}}\right) = 1 - \phi(1.25) = 1 - 0.8944 = 0.1056$$

2) Una máquina embotelladora puede regularse de tal manera que llene un promedio de μ onzas por botella. Se ha observado que la cantidad de contenido que suministra la máquina presenta una distribución normal con $\sigma = 1$ onza. De la producción de la máquina un cierto día, se obtiene una muestra de 9 botellas llenas (todas fueron llenadas con las mismas posiciones del control operativo) y se miden las onzas del contenido de cada una.

a) Determinar la probabilidad de que la media muestral se encuentre a lo más a 0.3 onzas de la media real μ para tales posiciones de control

b) ¿Cuántas observaciones deben incluirse en la muestra si se desea que la media muestral esté a lo más a 0.3 onzas de μ con una probabilidad de 0.95?

Solución:

a) Sean las variables aleatorias X_i : "contenido en onzas de la botella *i*" $i = 1,2,\ldots,9$

Entonces $X_i \sim N(\mu,1)$ para cada *i*.

Por lo tanto $\overline{X} \sim N\left(\mu, \frac{1}{9}\right)$. Se desea calcular

$$P(|\overline{X}-\mu| \le 0.3) = P(-0.3 \le \overline{X}-\mu \le 0.3) = P\left(-\frac{0.3}{\sigma/\sqrt{n}} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \frac{0.3}{\sigma/\sqrt{n}}\right) =$$

$$= P\left(-\frac{0.3}{\sigma/\sqrt{n}} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \frac{0.3}{\sigma/\sqrt{n}}\right) = P\left(-0.9 \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le 0.9\right) = \Phi(0.9) - \Phi(-0.9) =$$

$$= 2\Phi(0.9) - 1 = 0.6318$$

b) Ahora se pretende que

$$P(|\overline{X}-\mu| \le 0.3) = P(-0.3 \le \overline{X}-\mu \le 0.3) = 0.95$$

Entonces

$$P(|\overline{X}-\mu| \le 0.3) = P\left(\frac{-0.3}{\sigma/\sqrt{n}} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \frac{0.3}{\sigma/\sqrt{n}}\right) = P\left(-0.3\sqrt{n} \le \frac{\overline{X}-\mu}{1/\sqrt{n}} \le 0.3\sqrt{n}\right) = 0.95$$

Mediante la tabla de la acumulada de la normal estándar se tiene que

$$P\left(-0.3\sqrt{n} \le \frac{\overline{X}-\mu}{1/\sqrt{n}} \le 0.3\sqrt{n}\right) = 2\Phi\left(0.3\sqrt{n}\right) - 1 = 0.95 \;\Rightarrow\; \Phi\left(0.3\sqrt{n}\right) = 0.975 \;\Rightarrow\; \left(0.3\sqrt{n}\right) = 1.96$$

O sea $n \approx \left(\frac{1.96}{0.3}\right)^2 = 42.68$

Si tomamos $n = 43$, entonces $P(|\overline{X}-\mu| \le 0.3)$ será un poco mayor que 0.95

6.2 - Teorema central del límite

Acabamos de ver que la suma de un número finito *n* de variables aleatorias independientes que están normalmente distribuidas es una variable aleatoria también normalmente distribuida. Esta propiedad reproductiva no es exclusiva de la distribución normal. En efecto, por ejemplo, ya vimos que existen variables aleatorias discretas que la cumplen, es el caso de la Poisson y la Binomial.

En realidad, la propiedad que le da a la distribución normal el lugar privilegiado que ocupa entre todas las distribuciones es el hecho de que la suma de un número muy grande, rigurosamente un número infinito numerable, de variables aleatorias independientes con distribuciones ***arbitrarias*** (no necesariamente normales) es una variable aleatoria que tiene, aproximadamente, una distribución normal. Este es, esencialmente, el contenido del

Teorema central del límite (T.C.L.):
Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes con $E(X_i) = \mu$ y $V(X_i) = \sigma^2$ para todo $i = 1,2,\ldots,n$, es decir ***independientes idénticamente distribuidas***

Sea la v.a. $S_n = \sum_{i=1}^{n} X_i$ y sea $Z_n = \dfrac{S_n - n\mu}{\sqrt{n\sigma^2}}$.

Entonces $\lim_{n\to\infty} P(Z_n \le z) = \Phi(z)$, esto es $\lim_{n\to\infty} P\left(\dfrac{S_n - n\mu}{\sqrt{n\sigma^2}} \le z\right) = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx$

Dem.) sin demostración

Observaciones:

1- Notar que $E(S_n) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = n\mu$ y $V(S_n) = V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = n\sigma^2$

Por lo tanto $Z_n = \dfrac{S_n - n\mu}{\sqrt{n\sigma^2}}$ es la v.a. S_n ***estandarizada***

2- Notar que $P\left(\dfrac{S_n - n\mu}{\sqrt{n\sigma^2}} \le z\right) = P\left(\dfrac{\dfrac{S_n - n\mu}{n}}{\dfrac{\sqrt{n\sigma^2}}{n}} \le z\right) = P\left(\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}\right)$, por lo tanto también se puede enunciar el Teorema central del límite de la siguiente forma

Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes con $E(X_i) = \mu$ y $V(X_i) = \sigma^2$ para todo $i = 1,2,\ldots,n$, es decir ***independientes idénticamente distribuidas***

Sea la v.a. promedio muestral $\overline{X} = \dfrac{1}{n} \sum_{i=1}^{n} X_i$ y sea $Z_n = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$.

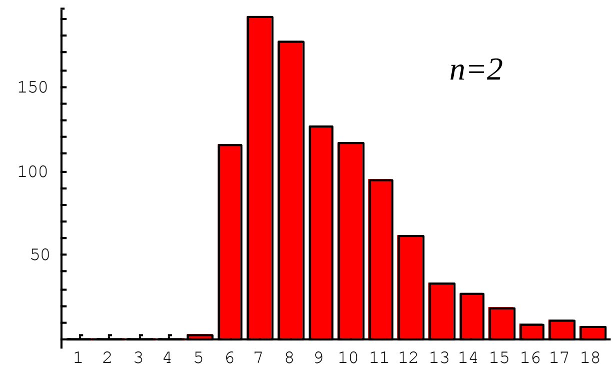
Entonces $\lim_{n\to\infty} P(Z_n \le z) = \Phi(z)$, esto es $\lim_{n\to\infty} P\left(\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \le z\right) = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx$

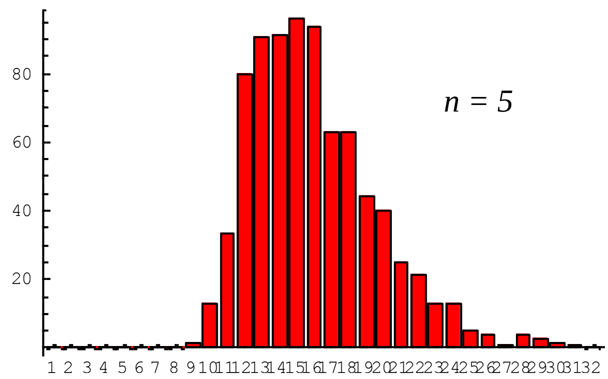
Donde $Z_n = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ es el ***promedio muestral estandarizado***

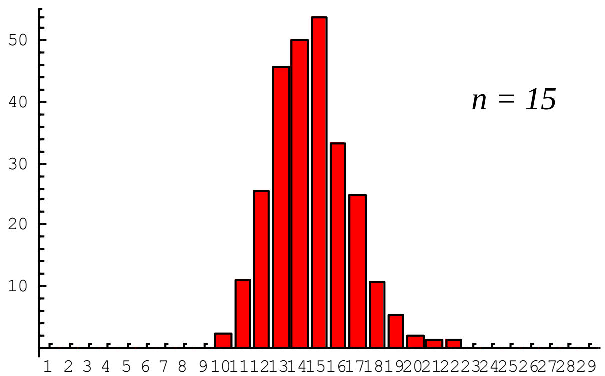
3- Aunque en muchos casos el T.C.L. funciona bien para valores de n pequeños , en particular donde la población es continua y simétrica, en otras situaciones se requieren valores de n mas grandes, dependiendo de la forma de la distribución de las X_i. En muchos casos de interés práctico, si $n \ge 30$, la aproximación normal será satisfactoria sin importar cómo sea la forma de la distribución de las X_i . Si $n < 30$, el T.C.L. funciona si la distribución de las X_i no está muy alejada de una distribución normal

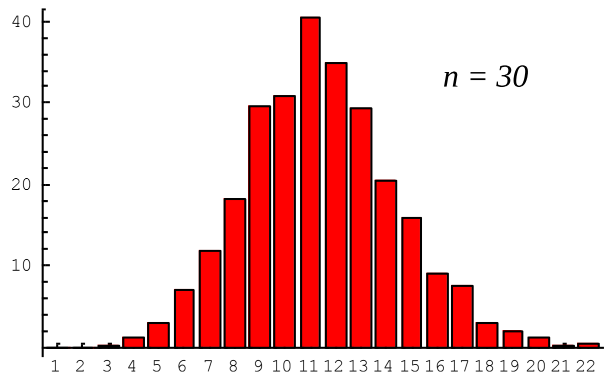
4- Para interpretar el significado del T.C.L., se generan (por computadora) n valores de una v.a. exponencial con parámetro $\lambda = 0.5$, y se calcula el promedio de esos n valores. Esto se repite 1000 veces, por lo tanto tenemos 1000 valores de la v.a. $\overline{X}$.

Hacemos un ***histograma de frecuencias*** de $\overline{X}$, esto es, tomamos un intervalo (a,b) donde "caen" todos los valores de $\overline{X}$, y lo subdividimos en intervalos mas chicos de igual longitud. La ***frecuencia de cada subintervalo*** es la cantidad de valores de $\overline{X}$ que caen en dicho subintervalo. Se grafican estas frecuencias obteniéndose los gráficos siguientes que se pueden considerar una aproximación a la verdadera distribución de $\overline{X}$.
Se observa que a medida que aumenta el valor de n los gráficos se van haciendo más simétricos, pareciéndose a la gráfica de una distribución normal.









Ejemplos:
1- Supóngase que 30 instrumentos electrónicos D_1, D_2, ...,D_{30}, se usan de la manera siguiente: tan pronto como D_1 falla empieza a actuar D_2. Cuando D_2 falla empieza a actuar D_3, etc. Supóngase que el tiempo de falla de D_i es una v.a. distribuida exponencialmente con parámetro $\lambda = 0.1$ por hora. Sea T el tiempo total de operación de los 30 instrumentos. ¿Cuál es la probabilidad de que T exceda 350 horas?

Solución:
Si X_i: "tiempo de falla del instrumento D_i" $i = 1,2,\ldots,30$

Entonces $X_i \sim Exp(0.1)$ para $i = 1,2,\ldots,30$

El tiempo total de operación de los 30 instrumentos es $T = \sum_{i=1}^{30} X_i$, donde

$$E(T) = E\left(\sum_{i=1}^{30} X_i\right) = 30 \times E(X_i) = 30 \times \frac{1}{0.1} = 300$$

$$V(T)=V\left(\sum_{i=1}^{30}X_i\right)=30\times V(X_i)=30\times\frac{1}{0.1^2}=3000$$

Entonces por T.C.L. $\dfrac{T-300}{\sqrt{3000}}\sim N(0,1)$ ***aproximadamente*** pues $n=30$

La probabilidad pedida es

$$P(T>350)=P\left(\frac{T-300}{\sqrt{3000}}>\frac{350-300}{\sqrt{3000}}\right)\underset{\text{T.C.L.}}{\approx}1-\Phi\left(\frac{350-300}{\sqrt{3000}}\right)=1-\Phi(0.9128)=1-0.81859=0.18141$$

2- Suponga que el consumo de calorías por día de una determinada persona es una v.a. con media 3000 calorías y desviación estándar de 230 calorías. ¿Cuál es la probabilidad de que el promedio de consumo de calorías diario de dicha persona en el siguiente año (365 días) sea entre 2959 y 3050?

<u>Solución</u>:
Definimos las variables aleatorias
X_i: "cantidad de calorías que una persona consume en el día *i*" $i=1,2,\dots,365$

Se sabe que $E(X_i)=3000$ y $V(X_i)=230^2$

Si $\overline{X}=\dfrac{1}{365}\sum_{i=1}^{365}X_i$ entonces $E(\overline{X})=3000$ y $V(\overline{X})=\dfrac{\sigma^2}{n}=\dfrac{230^2}{365}$

La probabilidad pedida es

$$P\left(2959\le\overline{X}\le3050\right)=P\left(\frac{2959-3000}{230/\sqrt{365}}\le\frac{\overline{X}-3000}{230/\sqrt{365}}\le\frac{3050-3000}{230/\sqrt{365}}\right)\underset{\text{T.C.L.}}{\approx}$$

$$\approx\Phi\left(\frac{3050-3000}{230/\sqrt{365}}\right)-\Phi\left(\frac{2959-3000}{230/\sqrt{365}}\right)=\Phi(4.15)-\Phi(-3.40)\approx1-0=1$$

<u>Aplicaciones del Teorema central del límite</u>

<u>Aproximación normal a la distribución binomial</u>

El Teorema central del límite se puede utilizar para aproximar las probabilidades de algunas variables aleatorias discretas cuando es difícil calcular las probabilidades exactas para valores grandes de los parámetros.

Supongamos que X tiene una distribución binomial con parámetros n y p. Para calcular $P(X\le k)$ debemos hacer la suma $P(X\le k)=\sum_{i=0}^{k}P(X=i)$ o recurrir a las tablas de la F.d.a. , pero para valores de n grandes no existen tablas, por lo tanto habría que hacer el cálculo en forma directa y muchas veces es laborioso.

Como una opción podemos considerar a X como suma de variables aleatorias más simples, específicamente, si definimos

$$X_i = \begin{cases} 1 & si \quad en \quad la \quad i - ésima \quad repetición \quad de \quad \varepsilon \quad ocurre \quad éxito \\ 0 & caso \quad contrario \end{cases} \qquad i = 1,2,\dots,n$$

entonces cada X_i se la puede considerar $B(1, p)$, y además $X_1, X_2, \dots, X_n$ son independientes

Podemos escribir $X = X_1 + X_2 + \dots + X_n = \sum_{i=1}^{n} X_i$ y ***si n es grande*** entonces X tendrá ***aproximadamente*** una distribución normal con parámetros np y $np(1-p)$, es decir

$$Z_n = \frac{X - n\mu}{\sqrt{n\sigma^2}} = \frac{X - n.p}{\sqrt{n.p(1-p)}} \approx N(0,1) \quad \text{si } n \text{ es lo suficientemente grande}$$

Observaciones:

1- La aproximación normal a la distribución binomial funciona bien aun cuando n no sea muy grande si p no está demasiado cerca de cero o de uno. En particular la aproximación normal a la binomial es buena si n es grande , $np > 5$ y $n(1-p) > 5$, ***pero es más efectivo aplicar esta aproximación cuando*** $np > 10$ ***y*** $n(1-p) > 10$

2- Corrección por continuidad.

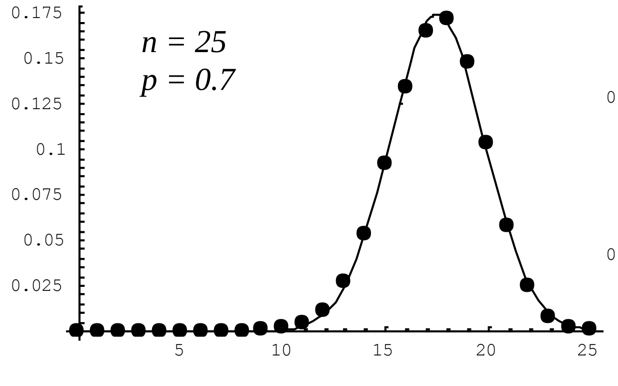
Acabamos de ver que si $X \sim B(n,p)$ entonces, para n suficientemente grande, podemos considerar que aproximadamente es $X \sim N[n.p, n.p(1-p)]$. El problema que surge de inmediato si deseo calcular, por ejemplo, la probabilidad de que $X = k$ (con k alguno de los valores posibles 0,1,2,…,n) es que la binomial es una distribución discreta y tiene sentido calcular probabilidades como $P(X = k)$ mientras que la normal es una distribución continua y, en consecuencia, $P(X = k) = 0$ puesto que para una variable aleatoria continua la probabilidad de que ésta tome un valor aislado es cero. Esto se resuelve si se considera $P(X = k) \approx P\left(k - \frac{1}{2} \le X \le k + \frac{1}{2}\right)$

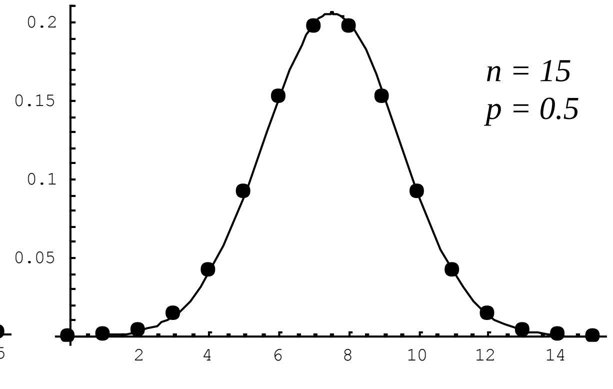
También se puede usar esta corrección para mejorar la aproximación en otros casos, específicamente en lugar de $P(X \le k)$ calculamos

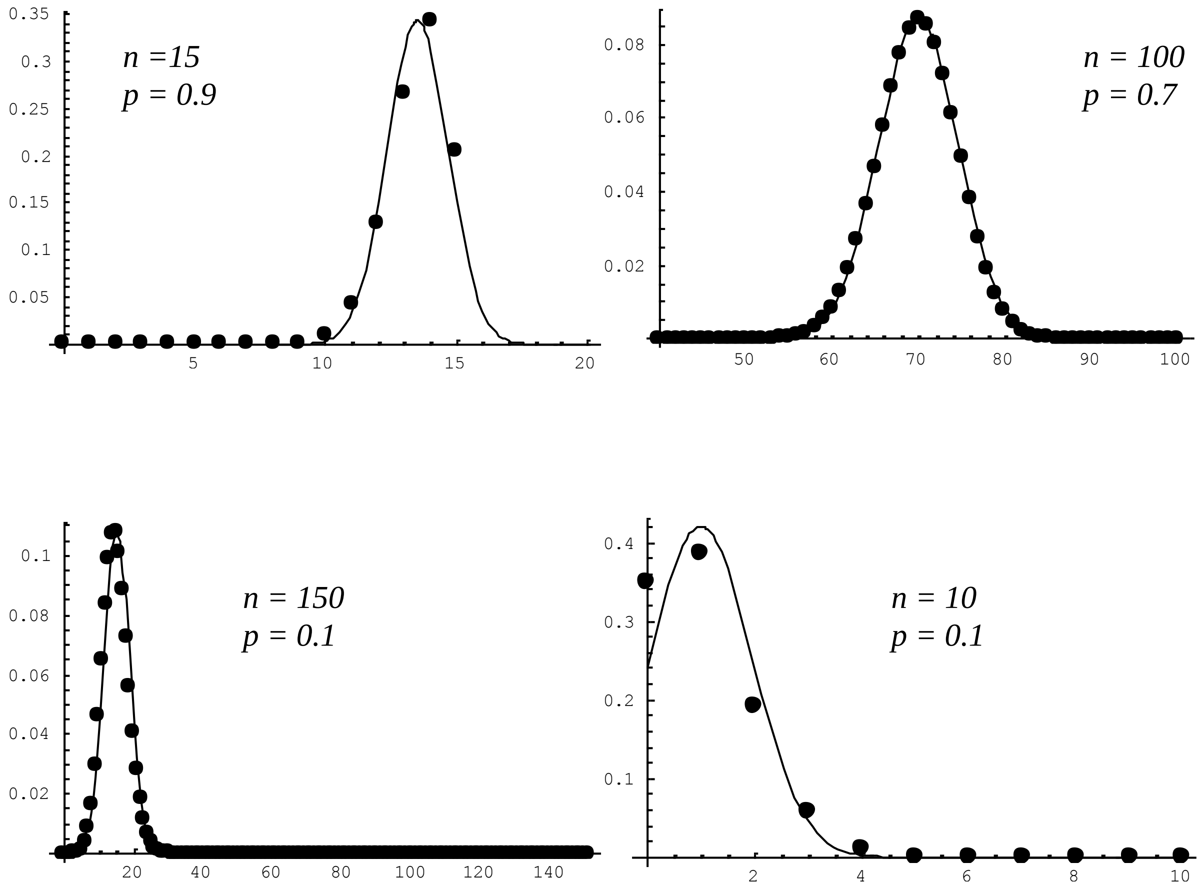
$$P(X \le k) \approx P\left(X \le k + \frac{1}{2}\right)$$

Y en lugar de $P(X \ge k) \approx P\left(X \ge k - \frac{1}{2}\right)$

En los gráficos siguientes se muestra para diferentes valores de n y p cómo aproxima la distribución $N(np,\ np(1-p))$ a la distribución $B(n, p)$







Ejemplos:

1- Sea $X \sim B(25,0.4)$. Hallar las probabilidades exactas de que $X \leq 8$ y $X = 8$ y comparar estos resultados con los valores correspondientes encontrados por la aproximación normal.

Solución:

De la tabla de la $F.d.a.$ de la binomial encontramos $P(X \leq 8) = 0.274$

Y $P(X = 8) = P(X \leq 8) - P(X \leq 7) = 0.274 - 0.154 = 0.120$

Ahora usamos la aproximación normal

$$P(X \leq 8) \approx P(X \leq 8.5) = P\left(\frac{X - np}{\sqrt{np(1-p)}} \leq \frac{8.5 - 10}{\sqrt{25 \times 0.4 \times 0.6}} \right) \approx \Phi(-0.61) = 0.2709$$

corrección por continuidad

Observar que el valor aproximado está muy cercano al valor exacto para $P(X \leq 8) = 0.274$

$$P(X = 8) \approx P(7.5 \leq X \leq 8.5) = P\left(\frac{7.5 - 10}{\sqrt{6}} \leq \frac{X - 10}{\sqrt{6}} \leq \frac{8.5 - 10}{\sqrt{6}} \right) = P\left(-1.02 \leq \frac{X - 10}{\sqrt{6}} \leq -0.61 \right) =$$

$$= 0.2709 - 0.1593 = 0.1170$$

Nuevamente este valor aproximado está muy cerca del valor real de $P(X = 8) = 0.120$

2- Suponga que el 10% de todos los ejes de acero producidos por cierto proceso están fuera de especificaciones, pero se pueden volver a trabajar (en lugar de tener que enviarlos a la chatarra). Considere una muestra aleatoria de 200 ejes y denote por X el número entre ellos que estén fuera de especificaciones y se puedan volver a trabajar. ¿Cuál es la probabilidad (aproximada) de que X sea

a) a lo sumo 30?
b) menos de 30?
c) entre 15 y 25 (inclusive)?

<u>Solución</u>:
Sea la v.a. X: "número de ejes fuera de especificaciones"
Entonces $X \sim B(200,\ 0.1)$, además $np = 200 \times 0.1 = 20 > 5$ y $n(1-p) = 200 \times (1-0.1) = 180 > 5$

Por lo tanto podemos aplicar la aproximación normal a la binomial

a) la probabilidad pedida es $P(X \le 30)$

$$P(X \le 30) \approx P(X \le 30.5) = P\left(\frac{X-np}{\sqrt{np(1-p)}} \le \frac{30.5-20}{\sqrt{18}}\right) \approx \Phi\left(\frac{30.5-20}{\sqrt{18}}\right) = \Phi(2.474) = 0.993244$$

b) La probabilidad pedida es $P(X < 30)$

Al ser X una v.a. ***discreta*** con distribución binomial $P(X < 30) = P(X \le 29)$

$$P(X \le 29) \approx P(X \le 29.5) \approx \Phi\left(\frac{29.5-20}{\sqrt{18}}\right) = \Phi(2.2391) = 0.98745$$

c)

$$P(15 \le X \le 25) \approx P(14.5 \le X \le 25.5) \approx \Phi\left(\frac{25.5-20}{\sqrt{18}}\right) - \Phi\left(\frac{14.5-20}{\sqrt{18}}\right) =$$

$$= \Phi(1.2963) - \Phi(-1.2963) = 2\Phi(1.2963) - 1 = 2 \times 0.90147 - 1 = 0.80294$$

3- El gerente de un supermercado desea recabar información sobre la proporción de clientes a los que no les agrada una nueva política respecto de la aceptación de cheques. ¿Cuántos clientes tendría que incluir en una muestra si desea que la fracción de la muestra se desvíe a lo mas en 0.15 de la verdadera fracción, con probabilidad de 0.98?.

<u>Solución</u>:
Sea X: "número de clientes a los que no les agrada la nueva política de aceptación de cheques"
Entonces $X \sim B(n,\ p)$ donde p es desconocido y es la ***verdadera proporción*** de clientes a los que no les agrada la nueva política de aceptación de cheques. El gerente tomará una muestra de n clientes para ***"estimar"*** p con $\overline{X} = \frac{X}{n}$ ya que $\overline{X} = \frac{X}{n}$ es la proporción de clientes a los que no les agrada la nueva política de aceptación de cheques ***en la muestra de n clientes***. Si no se toman a ***todos los clientes,*** entonces $\overline{X} = \frac{X}{n}$ ***no será igual a p***.

La pregunta es cuál debe ser n para que $\overline{X} = \frac{X}{n}$ se aleje del verdadero p en menos de 0.15 con probabilidad 0.98 por lo menos, o sea para que $P(|\overline{X} - p| \le 0.15) \ge 0.98$

Entonces planteamos

$$P(|\overline{X} - p| \le 0.15) = P(-0.15 \le \overline{X} - p \le 0.15) = P\left(\frac{-0.15n}{\sqrt{np(1-p)}} \le \frac{X-np}{\sqrt{np(1-p)}} \le \frac{0.15n}{\sqrt{np(1-p)}}\right) \approx$$

$$\approx \Phi\left(\frac{0.15n}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{-0.15n}{\sqrt{np(1-p)}}\right) = 2\Phi\left(\frac{0.15n}{\sqrt{np(1-p)}}\right) - 1 \geq 0.98$$

Por lo tanto $\Phi\left(\frac{0.15n}{\sqrt{np(1-p)}}\right) \geq \frac{0.98+1}{2} = 0.99$

Además $\frac{0.15n}{\sqrt{np(1-p)}} = \frac{0.15\sqrt{n}}{\sqrt{p(1-p)}} \geq \frac{0.15\sqrt{n}}{\sqrt{0.5(1-0.5)}} = 0.3\sqrt{n}$

Entonces debe cumplirse que $0.3\sqrt{n} \geq 2.33$ o sea $n \geq \left(\frac{2.33}{0.3}\right)^2 = 60.3211$

O sea ***se debe tomar una muestra de al menos 61 clientes***

<u>Aproximación normal a la distribución Poisson</u>

Se puede probar aplicando Teorema central del límite que

> Si $X \sim P(\lambda)$ entonces para λ suficientemente grande $\frac{X-\lambda}{\sqrt{\lambda}}$ tiene aproximadamente distribución $N(0,1)$

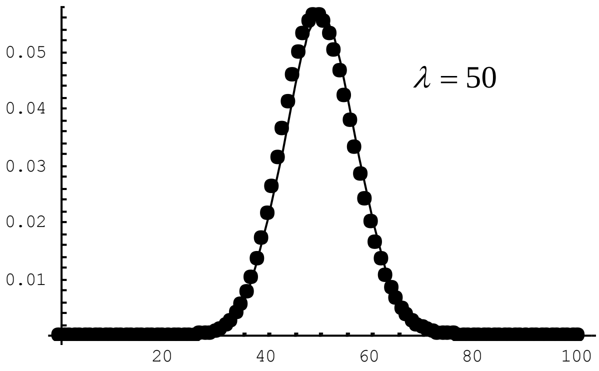
Es decir para λ suficientemente grande $\frac{X-\lambda}{\sqrt{\lambda}} \approx N(0,1)$
En la práctica si $\lambda \geq 30$ la aproximación es buena.

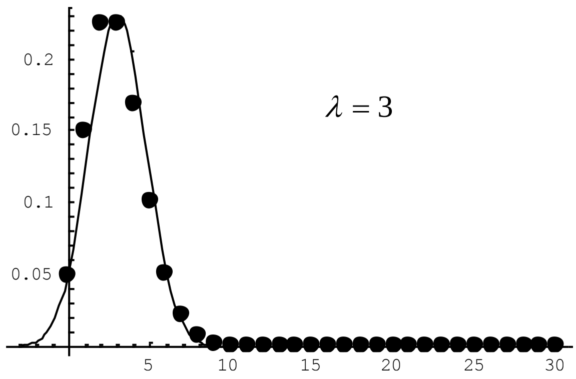
<u>Observación</u>: la demostración es sencilla si λ es igual a un número natural *n* pues, si consideramos las variables aleatorias $X_i \sim P(1)$ con $i = 1,2,\dots,n$ independientes, entonces ya sabemos que $\sum_{i=1}^{n} X_i \sim P\left(\sum_{i=1}^{n} 1\right)$, es decir $\sum_{i=1}^{n} X_i \sim P(n)$

Pero además por T.C.L. si *n* es grande $\sum_{i=1}^{n} X_i$ tiene aproximadamente distribución normal con parámetros $n\mu = n \times 1 = n$ y $n\sigma^2 = n \times 1 = n$

O sea la distribución de $\sum_{i=1}^{n} X_i$ que es exactamente Poisson con parámetro *n*, se puede aproximar con una $N(n,n)$, por lo tanto $\frac{X-n}{\sqrt{n}} \approx N(0,1)$ aproximadamente para valores de *n* suficientemente grandes

En los gráficos siguientes se muestra para diferentes valores de λ cómo aproxima la distribución $N(\lambda,\ \lambda)$ a la distribución $P(\lambda)$





Ejemplo:

El número de infracciones por estacionamiento en cierta ciudad en cualquier día hábil tiene una distribución de Poisson con parámetro λ = 50. ¿Cuál es la probabilidad aproximada de que:

a) más de 35 y a lo sumo 70 infracciones se expidan en un día en particular?

b) el número total de infracciones expedidas durante una semana de 5 días sea más 225 y a lo sumo 275?

Solución:

Sea X: "número de infracciones por estacionamiento en cierta ciudad en cualquier día hábil"

Entonces $X \sim P(\lambda)$ donde $\lambda = 50$

Como $\lambda = 50$ entonces $\dfrac{X-50}{\sqrt{50}} \approx N(0,1)$ (aproximadamente)

a) la probabilidad pedida es

$$P(35 < X \le 70) \approx \Phi\left(\frac{70-50}{\sqrt{50}}\right) - \Phi\left(\frac{35-50}{\sqrt{50}}\right) = \Phi(2.8284) - \Phi(-2.12132) =$$

$$= 0.997599 - 0.017 = 0.9805$$

b) Sea Y: "número total de infracciones expedidas durante una semana de 5 días"

Entonces $Y \sim P(\lambda)$ donde $\lambda = 50 \times 5 = 250$

La probabilidad pedida es

$$P(225 < Y \le 275) \approx \Phi\left(\frac{275-250}{\sqrt{250}}\right) - \Phi\left(\frac{225-250}{\sqrt{250}}\right) = \Phi(1.5811) - \Phi(-1.5811) =$$

$$= 2\Phi(1.5811) - 1 = 2 \times 0.94295 - 1 = 0.8859$$